

# Error Management in US Intelligence

## *Configuring Intelligence Processes to Reduce Error in End Products*

Andrew M. Shepard--CIA, 28 July, 1995

This paper is about planning for error in the collection, processing, analysis, and dissemination of intelligence. The focus is on common errors that affect the quality of intelligence--such as collection of inaccurate data and misinterpretation of raw reporting. We cannot prevent such errors, since all people and processes are imperfect, but we do have mechanisms to catch such errors and limit their impact. This paper highlights some major sources of error and the processes we use to detect and recover from them.

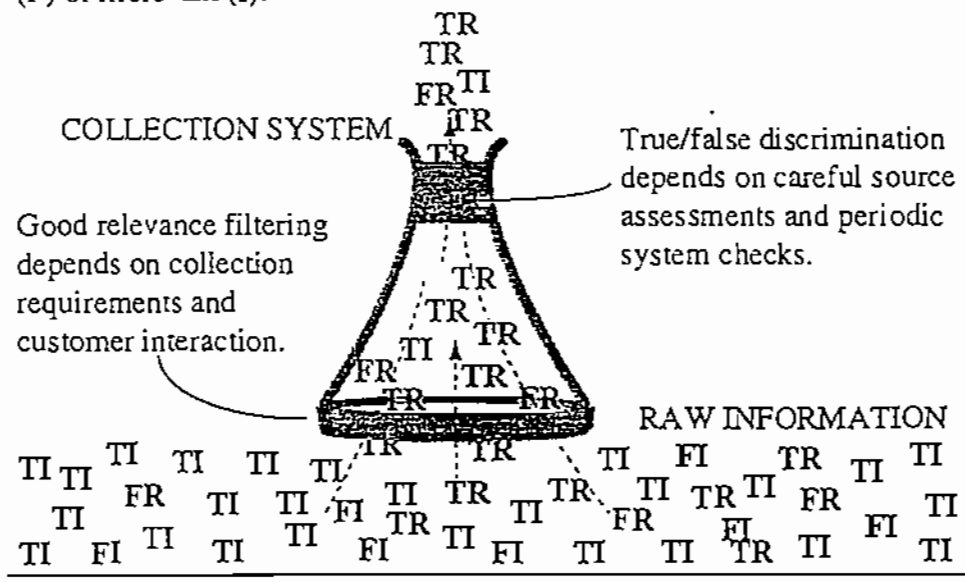
Errors are inherent to the way we choose to apportion the work of intelligence. US Intelligence is a collection of processes that, working together, approximate an ideal service. Error detection and recovery techniques are key to the process, helping to keep performance as close to the ideal as possible. But we can no longer improve or even sustain traditional techniques for error detection and recovery. To do much better, we need to revise our basic work processes--approximating the ideal in a new way that either generates fewer errors or allows more efficient detection and recovery. Our ideal has not changed, but our work reflects best practices as of the mid 1900s. Better options are available today because of modern information technology, which has fundamentally changed the best practices for organizing large and complex enterprises.

### The situation today--an overview

Major classes of error in the intelligence process relate to the stage where they occur--for example, in the original collection, during the processing of raw data, in the analysis of all-source data, or in transfers of information among these stages. Figure 1 depicts a generic collection operation, designed to find and report information that is accurate and important--information that is selected from a large amount of generally irrelevant information. The figure illustrates three kinds of error in collection:

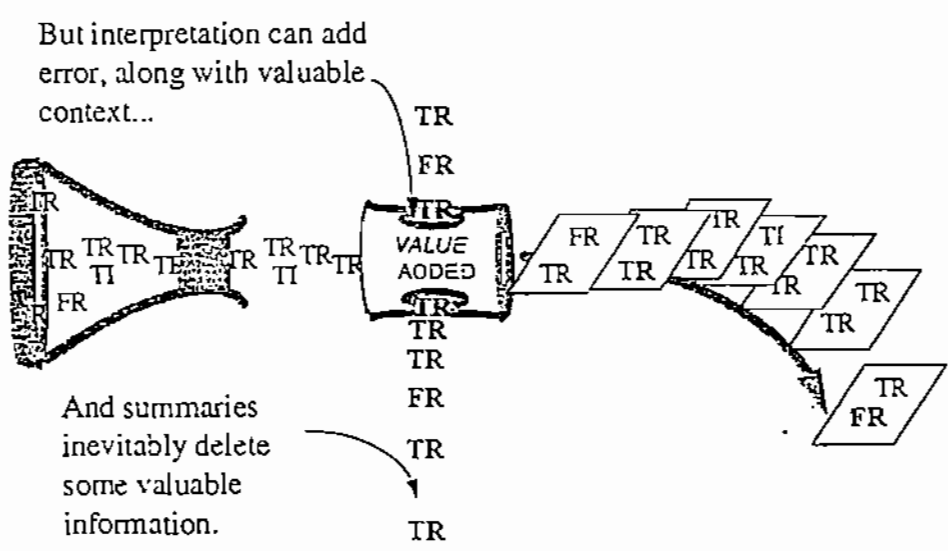
- Irrelevant information. Many collection operations yield a certain fraction of irrelevant material as an inevitable byproduct. Also, collection guidance may not be updated frequently enough, or may not be sufficiently detailed, to prevent collection of irrelevant information.
- False information. The collection sensor, transmitter, and receiver (or human equivalents in HUMINT) can all introduce error. Periodic checks on source reliability are used to guard against systematic causes of error.
- Overlooking relevant information. There are limits to the capacity of collectors; not all relevant information can be obtained. And, again, the collection guidance may not be adequately updated or sufficiently detailed.

Figure 1: Every collector looks for information that is true (T) and relevant (R). Occasionally, however, the information collected is false (F) or irrelevant (I):



As illustrated in Figure 1, collectors have mechanisms to help prevent errors. In addition, collectors generally catch some errors through an internal process of review, as shown in Figure 2. This processing does a variety of things--removing redundant information, consolidating related information into a coherent report, adding background information that provides context, and detecting irrelevant or false information. This value added processing also helps to spot gaps in the information and suggest additional collection--helping to reduce errors of omission.

Figure 2: Most collectors add value by processing, summarizing, interpreting, and packaging the collected information...

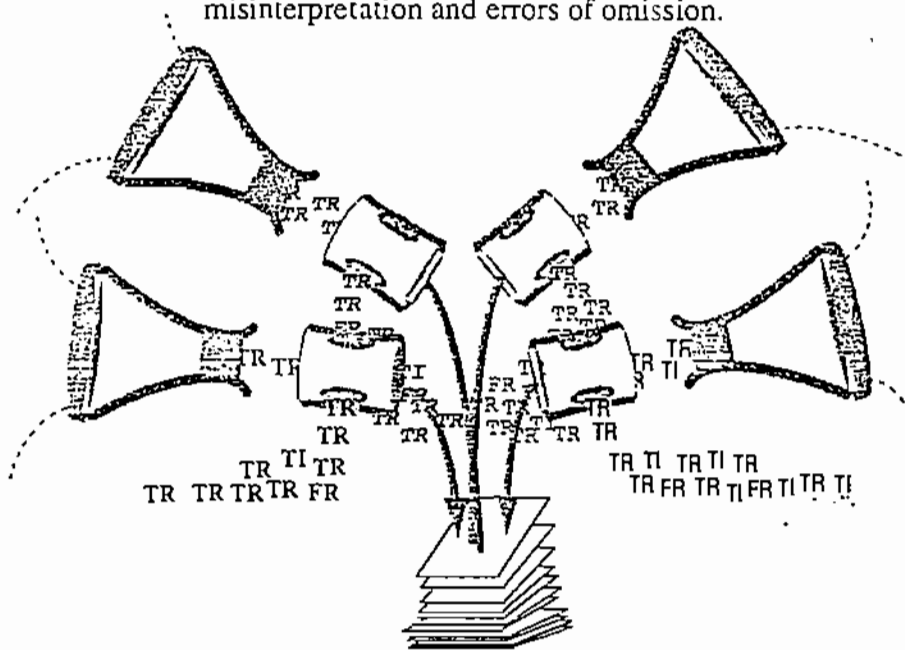


The process of adding value to collected information can introduce new errors, even as it helps detect and correct errors of collection. Under pressure to remove irrelevant information and to consolidate a large amount of information into a terse report, employees tend to set aside information that their supervisors would not recognize as important, as illustrated in Figure 2. However, since the true relevance of any new piece of information can depend on its correlation with all-source data that is not available to the collector, this selection process tends to sidetrack a large amount of potential useful data. Providing context and interpretation during the value added processing also can be a source of error. In order to ensure that readers understand the implications of a new piece of information, collectors put the new data into context or into a more readily exploitable form. Some amount of error is a price we pay to enhance the overall impact of the data.

Error can also be caused, even when the right information is selected and reported accurately, if the process of adding value significantly delays the release of good information. All the users downstream who correlate information from multiple sources, to identify important developments or to target collection assets, make decisions each day based on whatever "pieces of the puzzle" are available to them. Delays in reporting can deny a key piece of information or prevent the user from recognizing the importance of other data that he or she has in hand.

Downstream, the users of reported information often are so inundated with reports that they cannot afford to look at anything that is not immediately recognizable in its relevance and importance. They rarely clamor for more of the information that is being shed by collectors in the reporting process today. As suggested in Figures 3 and 4, most users receive reports from multiple sources, and have more than enough to read.

Figure 3: The large volume of data from all sources inhibits effective use by collection and processing components--contributing to misinterpretation and errors of omission.



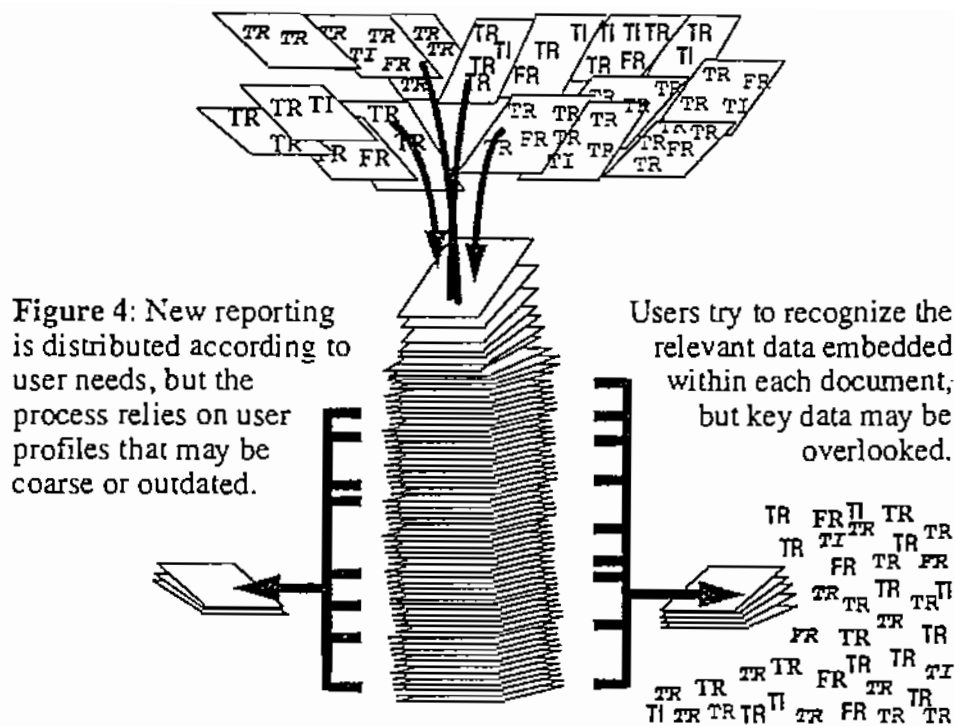


Figure 4 reflects our current process for distributing new information to all-source users, as documents that are matched against each users' needs. We are familiar with several kinds of error in this process:

- The matching of documents to user needs is imperfect. Technology for filtering documents is imperfect, to begin with--but users who depend upon it also lack the full skills (or time) required to make their queries and interest profiles reflect their real needs. They miss some valuable documents and receive many documents that are not relevant to their needs.
- The process of reading documents introduces error, especially when done quickly. The overall volume of material leads the user to rush through relevant data in order to avoid missing any crucial information in the remaining documents. In this rushed process, some relevant information goes by unseen. Also, the daily mix of relevant and irrelevant documents leads to a habit of deciding very quickly whether a document warrants close inspection--and users sometimes dismiss as irrelevant a document that has potentially valuable information.

Errors in reading are not highly visible, since they tend to be errors of omission that may never be discovered. From a distance, all we see is that collected information is being delivered, users are busy reviewing documents, and people are obtaining valuable information. The extent to which some information is not getting to users when it should, or is not being seen by users when it should, is not easily measured. We know about the problem primarily because of unexpected discoveries of data, found by analysts who do

retrospective searches. The extent of the problem varies with the intelligence topic.

Although we cannot measure the extent to which data is being missed, we can make a rough estimate. Several studies of the technology we use for search and retrieval and for document distribution suggest that users, on average, will miss about 25% of the relevant documents that are available. If we assume that 5% of the delivered documents are also erroneously overlooked as irrelevant, then we can estimate that users see about 70% of the available documents generated by collectors. We can also assume that users are successfully spotting, reading, and remembering most of the relevant material within the documents they do choose to read--say, 80%. If this is anywhere near the true situation, our analysts, target planners and others are exploiting a little more than half of the relevant information that collectors generate for use in meeting consumer needs.

Another source of error--in the form of information loss--occurs when users do not capture critical information for later retrieval and use. In the act of reading and correlating information from many documents, the user derives insights that exist only in that reader's memory, unless some action is taken. The simplest action is to file the document, in a topic folder or project folder, as illustrated in Figure 5. (Topic folders ensure that the document can easily be retrieved when searching for all relevant material on that topic--so

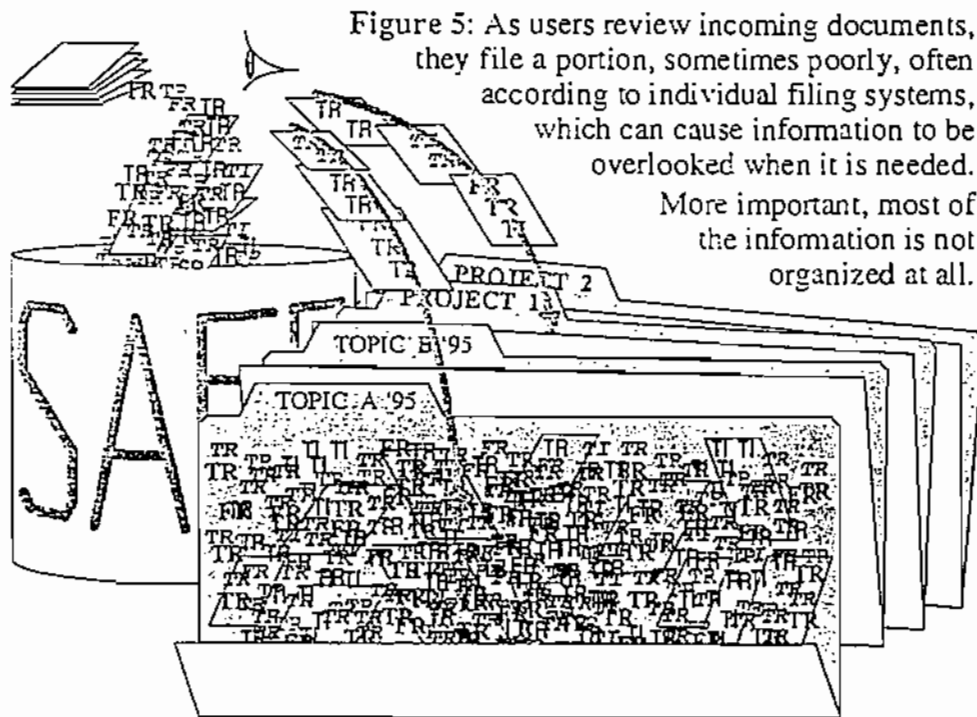


Figure 5: As users review incoming documents, they file a portion, sometimes poorly, often according to individual filing systems, which can cause information to be overlooked when it is needed. More important, most of the information is not organized at all.

\* One implication of this estimate of missed data is that a marginal dollar of intelligence resources, spent on additional collection, might only yield about 50 cents of additional information that is actually used. The same dollar--invested so as to increase the percentages mentioned above--might result in many dollars' worth of additional information finding its way into our support to intelligence consumers.

that the user can save time and avoid repeating the 25% loss mentioned above. Project folders are used to assemble documents for a specific task, after which the file may be saved or not.) Aside from filing, the user might summarize information in a time line, extract pieces of information to be cited later, or enter pieces of information into a relational database. These actions occur only rarely, however, because of time constraints.

As shown in Figure 5, most of the incoming information that is reviewed today is not saved in any way except in very large holdings of documents (for example, SAFE). Of the information that does get saved in folders today, much of it is saved in personal files that do not conform to any institutional filing scheme. As a result, it may be hard for successors to take advantage of their predecessors' folders--and colleagues are rarely able to benefit from the user's files.)

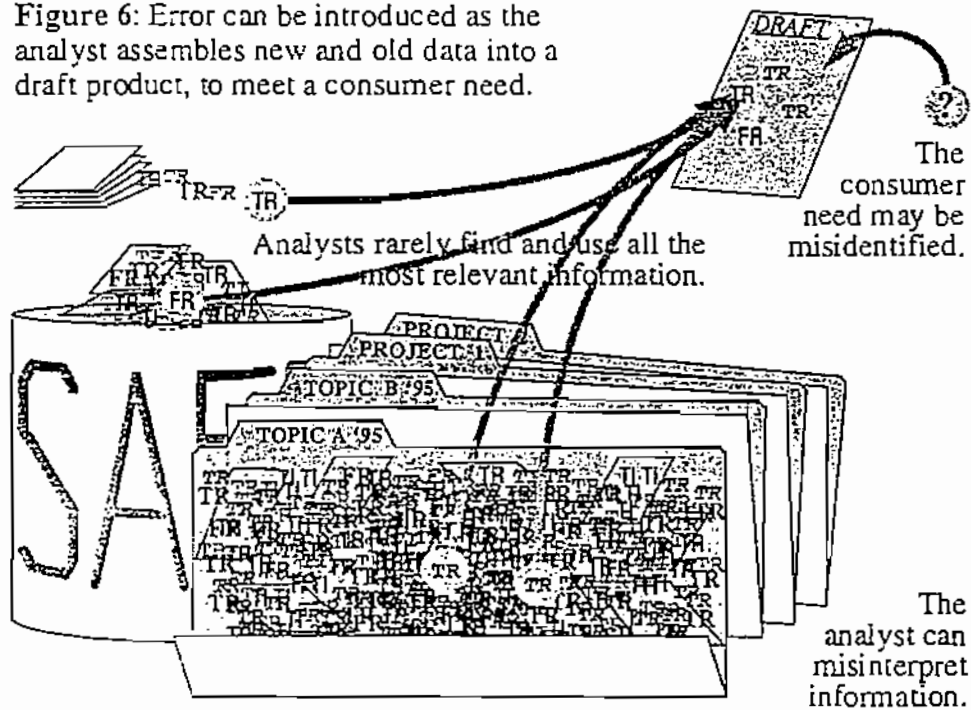
For some kinds of intelligence questions, files can often support a quick response. For example, if a consumer asks whether the Urbanians are smuggling frambits aboard any European-flagged vessels, the analyst can provide an affirmative response as soon as he or she comes across instances in the file. If, on the other hand, a consumer asks for a list of countries whose carriers have been used in the last two years to smuggle frambits into Urbania, we would be lucky to find an analyst who had filed the necessary information to formulate a comprehensive list without extensive work. Most of the questions we handle, of this type, are not predictable enough to warrant maintaining a comprehensive file on the subject for years in advance. Therefore, most topic files tend to be fairly general.

Because topic files tend to be quite general, the user often has difficulty deciding which general topic folder is the best one in which to save a newly arriving document. A three-page document on Urbanian defense spending and modernization plans might include one paragraph mentioning frambits; it might very sensibly be filed in a folder entitled "Urbanian Defense Mod," or "Urbania-'96 budget plan." When the question arrives about Urbanian frambit smuggling, the analyst typically will have to search several folders that contain the kinds of documents that might include relevant information. Because of the generality of file topics, therefore, we sometimes make errors of omission even when users have taken the time to file information. Because some of the relevant information may not get filed--or may not get filed correctly, or may not have been filed in the folder that the user thinks to look in--the information that is actually considered in answering a consumer question may well be incomplete.

There are additional errors that can occur in the course of addressing a consumer need, as illustrated in Figure 6. Perhaps the most critical source of error is misunderstanding about the needs of the consumer. Information about the consumer need affects all the many decisions that must be made as to what data to provide, and in what context, when generating a briefing or paper. A brief communication between consumer and intermediary can leave considerable room for mistaken judgments throughout the production and review hierarchy, concerning the information and context most likely to help. Also, as noted above, the analysts responsible for addressing a particular need can fail to find a significant fraction of the available information--including data that might be

needed directly or data that might change the user's interpretation of information that is in hand. Finally, for this reason or many others, the analyst can simply misinterpret whatever information is reviewed for the consumer, in the time available.

Figure 6: Error can be introduced as the analyst assembles new and old data into a draft product, to meet a consumer need.



Our primary mechanism for catching and mitigating errors in the final stages of intelligence support to consumers is the coordination process. Coordination among multiple analysts allows more information to be brought to bear, allows faulty data or interpretations to be discovered, and also affords a chance for more information about the consumer to be taken into consideration--because each coordinating officer brings different expertise, files of information, and knowledge of the consumer to bear in reviewing the product. Management review performs a similar role--but without their own files or familiarity with the raw data, managers primarily add value by examining the logic and clarity of whatever information is contained in the draft. Managers also bring general knowledge of the consumer to bear, and one or more areas of special expertise. Finally, managers help resolve differences that arise in the coordination process (some of which may be bureaucratic in nature, inhibiting the intended benefits of peer review).

Much has been written and said about the pros and cons of the coordination and review processes in intelligence. For this paper, it is sufficient to note that either peer coordination or management review can introduce error as well as catch error. The hope and belief is that both do more to reduce error than to increase it. Nevertheless, several

trends and factors suggest that we cannot rely on coordination and management review to work as well in the years ahead:

- The need to reduce staff levels. This calls into question the relative cost and benefit of each set of hands that touches a draft.
- A trend toward shorter deadlines, as we strive to be relevant to the immediate needs of consumers, and as global commercial news and communications tend to bring new information to consumers on a shorter cycle.
- A general increase in the volume of raw information that should be monitored and considered--leaving less time for the participants in a peer review process to perform their roles as thoroughly or thoughtfully.
- A trend toward more complex multi-country, multidisciplinary issues that consumers are facing in a world that is multipolar, economically more integrated, and increasingly affected by nongovernmental parties. This calls for numerous specialists to be involved in addressing a typical consumer need.
- An overall increase in the number of actions going on, as we shift operations to meet the specific needs of key individual consumers, rather than trying to serve them all with fewer, more general products. The time that can be devoted to coordination and review of a larger number of products is necessarily reduced.

### The situation could be quite different

If the architects of a new intelligence service were starting from scratch today, they would think differently about options for limiting error in the overall process of collecting and analyzing information to meet consumer needs. Two widely recognized principles of organizing information work, not applicable in 1947, would be central to the design:

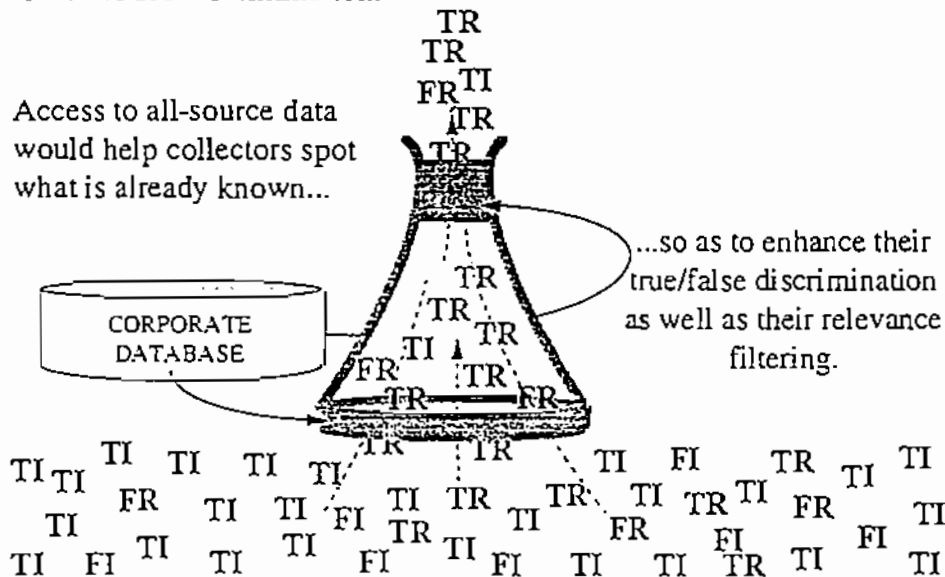
Because information work takes place on electronic systems, work can be organized and managed regardless of the geographic location or bureaucratic structure in which the participants find themselves.

Because information can be disseminated and stored in widely accessible systems, any employee who needs information to make a decision can get it directly.

These two principles of work in the "information age" replace many of the premises upon which our current methods and structure are based. In the early years, it was necessary to divide the labor of intelligence into discrete and largely self-managed elements for each type of collection, and for each type of analysis. Managing the process of "deciding what to collect," for example, necessitated elaborate mechanisms for requirements to filter up from many analytic components, and for these requirements to be reviewed, distilled, prioritized, and distributed to many collection components--there to be disseminated to appropriate collection officers. Manual processes required a hierarchical approach. Seasoned officers and extensive clerical support were required throughout the decision making process.

Today, computers have done away with much of the need for clerical support, but the process still entails reviews and approvals by seasons officers on behalf of ultimate actors. Designing a process from scratch in the 1990s, one could take a more streamlined approach. Corporate information about what we know, and need to know, could be organized for direct access by all who have a need to know, including collectors. Access to corporate holdings could reduce error in collection and reporting, as illustrated in Figures 7 and 8.

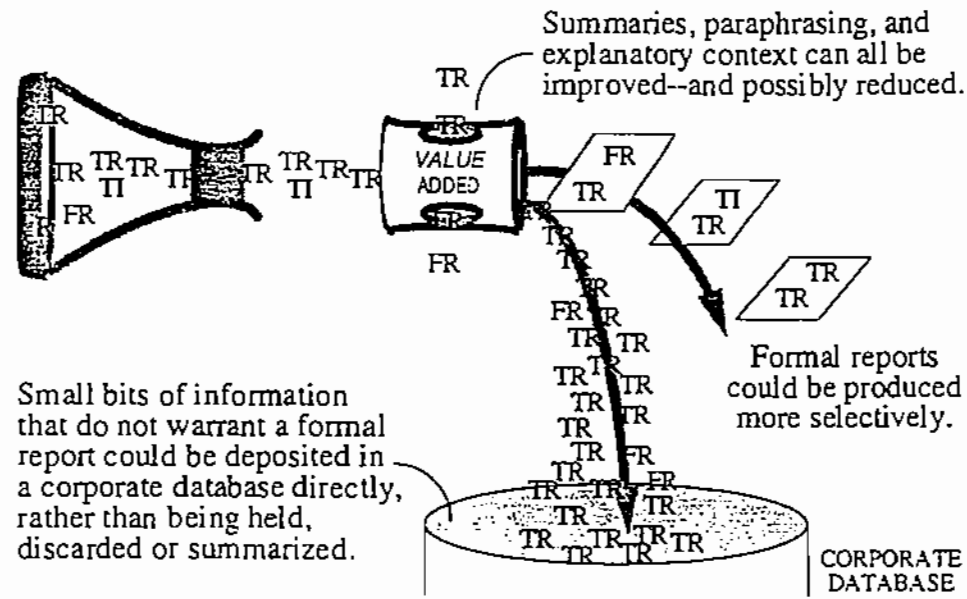
Figure 7: Collection of accurate and relevant information could be enhanced by improving the information used for relevance filtering and for true/false discrimination.



Modern intelligence architects could organize intelligence on foreign people, organizations, and events in a corporate database that allows easy correlation of information from many different documents and sources. As suggested in Figures 7 and 8, the creation and use of such a database could significantly aid collectors in their filtering and reporting. In particular, much of the collected information could be disseminated directly to the corporate database, without having to be assembled into a coherent, single-source, policy-relevant "story" that warrants a formal single-source report. The simple fact that Person X works for Organization Y could be added to the corporate knowledge, for example--information that would not be deemed report-worthy today, even though it might be a key link in assessing foreign plans or transactions. The simplicity of direct

\* The term "corporate information," as used in the external literature, generally refers to data needed by employees throughout the corporate enterprise to do their work. It includes a lot of internal proprietary information, but not personnel files, for example. It includes archival information and drafts of work in progress, but not personal notes and working papers. In the postulated intelligence context, corporate holdings probably would include what we know about consumer needs, and what we know about foreign people, organizations, etc--but would not include operational information or employee working files.

Figure 8: Errors of omission and misinterpretation can be reduced if the people and systems adding value can access corporate knowledge.



entry into the corporate database could make dissemination of such tidbits affordable for the first time. Labor intensive drafting and review of formal reports could be reserved for more complex and important news.

The presence of a corporate database would also enable designers of a modern intelligence service to reduce the information overload that constrains users today. Collectors and analysts who need information organized by country, topic, person, organization, etc. would find much of it already organized--because it would enter the corporate database in an organized fashion when first disseminated. Moreover, a service of common concern could be designed to extract key information elements from a wide range of documents for entry into the corporate database, on behalf of all users (discussed separately).<sup>\*</sup> Therefore, although collectors would disseminate more data, the increased volume would not be apt to aggravate user errors caused by information overload today. Instead, user errors relating to imperfect document dissemination, failure to recognize relevant information, and imperfect filing and lack of organized holdings could all be significantly reduced.

<sup>\*</sup> For a discussion of some conceptual options for affordably organizing incoming information from documents into a corporate database, see "Keeping a Watchful Eye on Incoming Information, At Less Cost--A Process Innovation to Cue Employees to Important New Developments," 7 June 1995.

Catching errors in the use of information to meet a specific consumer need would continue to depend on bringing multiple people's expertise to bear on draft products. In a modern design for the intelligence process, the opportunity to catch errors would be enhanced by two factors:

- Errors of fact in the raw data would be more likely to show up as inconsistencies with all the other data reported. Because much more all-source data would be available for corroboration, clarification, or refutation, all-source analysts would be less likely to use erroneous or outdated information.
- Coordination and collaboration would be greatly simplified--by common access to data, easy attachment of background data to draft products, and the ability to coordinate or collaborate easily with employees who are outside the traditional concept of "producers."

Recalling the principles of modern information work, any intelligence product generated for an important consumer could benefit from the involvement of employees located anywhere in the intelligence enterprise--whether a collector in the field, an analyst on rotation downtown, or a collection manager who has interacted extensively with the consumer in question. All can be part of a response team that operates over the corporate network to meet the specific need in question. One of the key benefits from this flexibility is that a smaller intelligence organization can still bring a wide range of expertise and experience to bear on whatever task is most important to the mission. In terms of managing error, this flexibility means always being able to involve the people who know most about the consumer need, the quality of the key source data, and the analytic assumptions. These three areas of knowledge, together with access to all the pertinent data, make up the factors most likely to cause or catch errors of the sort that have been discussed in this paper.

## Conclusion

Errors are made every day throughout the intelligence process--errors of the sort that are nobody's fault but which are simply the result of having to divide the work of intelligence among a very large number of people--an inherently error prone situation. We recognize the potential for error and invest considerable resources in prevention and mitigation techniques. But we could do much better--both in reducing error and in reducing the cost of managing error--if we organize intelligence collection, analysis, and consumer support around the concept of corporate communications and data holdings.

This paper is meant only to stimulate the reader's thinking--not to prescribe any specific new architecture for intelligence workflows. It is intended to highlight the concept of inherent error, and error management, as a factor in the ongoing discussion of intelligence reforms and work process options.

---

\* For a discussion of options for applying corporate expertise more effectively, see "Meeting the Needs of Intelligence Consumers--A Corporate Process to Serve More People More Fully," 13 June, 1995.