TEACHING THE GIANT TO DANCE: CONTRADICTIONS AND OPPORTUNITIES IN OPEN SOURCE WITHIN THE INTELLIGENCE COMMUNITY

BILL STUDEMAN

It's a pleasure to be here today to take part in this symposium. This is truly an exciting time -- a revolutionary time -- to be in the intelligence business. And no area is full of more promise for intelligence than open source access and exploitation. In my only reference to dancing, you can consider this presentation as an attempt to describe how we move the Intelligence Community giant from the square dancing era of yesterday to the open source Lambada of tomorrow.

A number of people that I've talked to -- including Members of Congress, journalists, and the public -- have asked me to explain why intelligence organizations are interested in unclassified information. So I'd like to begin by asking a rhetorical question: "Why does the Intelligence Community collect and analyze open source data?"

This is not a new issue for intelligence. As you know, intelligence has drawn broadly on open sources for many years. FBIS -- the Foreign Broadcast Information Service -- has collected, analyzed and reported open source intelligence from all over the globe for over half a century -- and it has done a superb and highly valued job.

The information that FBIS has collected over the years has been critical to US national security decisionmakers.

- Few people realize that during the 1962 Cuban Missile Crisis, through monitoring Radio Moscow, FBIS provided President Kennedy with the first news of the Soviet decision to withdraw missiles from Cuba.

- At times, public information has been the Community's only source during a crisis. For example, during the 1956 Hungarian uprising and the 1968 invasion of Czechoslovakia, radio broadcasts provided intelligence analysts with highly relevant, timely understanding of what was happening in the streets of Budapest and Prague.

- Open sources can also provide tip-off or early warning of events. You may be interested to know that the change in policy leading to the February 1989 Soviet military withdrawal from Afghanistan was identified in FBIS media analysis reports as early as May 1985.

- More recently, CIA analysts who monitored the collapse of the Soviet Union estimate that at least 80% of their information came from open sources.

• And now, intelligence analysts are using open sources to monitor the Yugoslav crisis. They study transcripts of radio broadcasts by both sides to gain insights into the intensity of the conflict and the probable outcome.

Although all of the Intelligence Community has relied on open sources for many years, some people who are not familiar with our business are confused about how intelligence uses information that is available to the public.

The first thing that people must understand is that intelligence is not competing with the media. But intelligence and the media are in the same business; that is, ultimately, to tell a story of relevant interest, but in our case, the story normally relates to a threat or a foreign issue of high or potential interest to U.S. or allied policymakers, planners, or warfighters. Our goal is not necessarily to produce raw open source data, but to glean information from open sources that is of interest to intelligence as background reference material for collectors and analysts/producers, and, more importantly as a source of information to be fused with data from classified sources and methods and this is again principally for the government customer.

Our analysts rely on a multiplicity of sources -- including signals, imagery, human and other classified intelligence sources as well as openly available data -- to produce their reports.

Foreign intelligence and counterintelligence earns its money first by maximizing these classified sources and methods, and secondly by building highly structured analysis and production systems which are highly responsive to the widest range of U.S. and allied customers, be the topic political, military, economic, environmental, sociological, law enforcement support, or otherwise.

But good intelligence officers, like media personnel, are essentially information hounds. The highest emphasis is placed on timeliness, relevancy, accuracy, and completeness of data disseminated at the lowest and most readily usable classification level and tailored to the diverse sets of simultaneous users at varying echelons of the bureaucratic structure from the President to lowest platoon leader and beyond.

The highest form of intelligence enlightenment is the dynamic and continuous fusion of data from all available sources. In this blending process a great synergy results, and this magic cannot be accomplished without unconstrained and continuous access to open source data. Open source can

provide event specifics, background context, focus, contrast, improved accuracy, alarms, and many other positive features associated with data manipulation in an information age.

While untrustworthy data can often be associated with classified sources and methods, open source data can be a frequent source of biased and misleading information, or worse yet, the product of deliberate deception or information control practiced in parts of the world by a less free press that may also operate as a propaganda instrument of government forces. This dictates that a strong data evaluation system be in place for use with open source data, as it is for classified data.

On the positive side, when an open source contradicts other intelligence sources -- or other open source reporting -- it serves as a flag for the analyst to re-evaluate his or her analysis. For example, at a time when there was wide intelligence speculation that the Dominican Republic might extradite a terrorist, an FBIS report called attention to a press account that the Dominican president had said he would not extradite the terrorist.

Utilizing intelligence analysis techniques, it is frequently possible to interpret or predict events based on open source usage. The evidence is often acquired through laborious textual analysis -- and by comparing media content with past actions.

- For example, FBIS analysts anticipated the February 1979 Chinese invasion of Vietnam by demonstrating that, with rare exceptions, the wording of authoritative Chinese warnings to Vietnam had only been used in instances in which Beijing had actually applied military force.

On some occasions, intelligence analysts adjudge open source information to be more accurate than classified sources. This can derive from either the weight and credibility of open sources versus the untested, contradictory or poor performing nature of the classified sources or from some other evaluation criteria. When we do favor unclassified sources over classified sources, we need to be sensitive to the credibility we have lent the data by adopting it as our own position.

Most people in our business agree that open sources have proven to be enormously invaluable to intelligence. Even during the Cold War, when intelligence was focused principally on acquiring secret information, open sources gave us some highly usable glimpses into closed societies. Today, with a generally more open world and a considerably more free and independent world press, open sources have even

greater value for intelligence. In the new global environment, open sources provide much more hard, credible data about a wide range of international political, social, and economic issues.

There is a complex relationship between the way open source material is mixed with classified data and the concept of openness. We frequently have products where only a small amount of the overall data comes from classified sources requiring security protection. We have security procedures in effect to clearly mark paragraphs which possess classified data, and thus enables much greater sanitization of intelligence publications to the unclassified level. The more complete and expansive the open sources, the more likely we can produce a wider variety of unclassified or lower classification products using the classified data as background for confidence-building and credibility. It is important to recognize that once an Intelligence Community agency puts its name on an essentially unclassified product, it may assume an enhanced credibility beyond that of the original open sources. This obligates the Intelligence Community to high standards of quality control, which we would expect of our people, in any case.

The Intelligence Community's current challenge is to expand the use of open sources to cover a broader range of issues -- such as weapons proliferation, economic competitiveness, and the environment. As one example, it is estimated that some 80% of the information needs for <u>environmental</u> intelligence can be met through information that is available to the public.

As you are well aware, the quality and quantity of open source information continues to grow:

- We have identified some 8,000 commercial data bases -- and the vast majority have potential intelligence value.

- The number of worldwide periodicals has grown from 70,000 in 1972 to 116,000 last year.

- The explosion of open source information is most apparent in the Commonwealth of Independent States, where <u>today, there are</u> some 1,700 newspapers that were not published three years ago.

- While the number of TV and radio stations around the world has not experienced rapid growth, the broadcast time and breadth and depth of their coverage, and the availability of cable TV are clearly on the upswing.

• The sources of "gray literature," (i.e., private or public symposia proceedings, and academic studies) around the world are also increasing dramatically.

As you know, open source encompasses a wide array of mediums -- including printed material, such as books, magazines and newspapers; as well as maps, photographs, data files, digital imagery and broadcast media -- both radio and television. These multimedia open source inputs correspond well to the range of product outputs used by the Intelligence Community. These intelligence outputs tend to be multimedia in nature, including hard copy and electronic dissemination of written and formatted, man and machine readable text, imagery, graphics, maps, and other situational displays as well as video. Because of the need to move data quickly in worst case situations, electronic information handling and display systems are most common. Most of these intelligence systems use off the shelf hardware, open systems, commercial achitecture and operate as part of large area networks. The Intelligence Community has a special problem managing multilevel security in systems where open source data is mixed with classified data.

I'd like to say a special word about TV -- because it is a relatively new area for intelligence. Each week, FBIS monitors 790 hours of television from over 50 countries in 29 languages. Foreign TV programs -- such as news programs and documentaries -- give analysts a multidimensional feel for a country or material that other open source media cannot provide. Many analysts prefer to see the way a particular country chooses to portray events visually, rather than relying on the news network "filter." Coverage of foreign television brings us closer to what is happening in all areas of the world; it allows us to monitor crises as well as to broaden our knowledge of more restrictive societies. For example, the revolutions in Eastern Europe were covered extensively on those countries' domestic television.

In addition to analyzing foreign television, intelligence organizations are producing classified videos for policymakers which incorporate information from foreign news programs. The end result is a high-impact intelligence product used exclusively in the government that improves policymakers' understanding of complex issues.

The dramatic increase in open source material, its wide variety -- and its increasing value to intelligence -- demand a revolutionary change in the Intelligence Community's approach to open source management, collection, processing and dissemination.

Unlike the other collection disciplines, which are highly structured, open source is not a tightly integrated discipline in the Intelligence Community. Over the years, open source information collectors, processors, and users have been diverse and decentralized groups spread across the breadth and depth of the Community. As a consequence, the various agencies in the Community didn't know the extent of unclassified holdings of other agencies, and had virtually no capability to share electronically the information which they did possess.

In short, the Community lacked a unifying structured, and a coherent and consistent set of overall requirements for the collection, processing, exploitation, and dissemination of open source information.

A DCI Task Force was formed last year to make recommendations on these issues -- and important changes are underway. As a result of the task force, for the first time the DCI has established an Open Source Coordinator [Paul Wallner of the Defense Intelligence Agency] who is:

- cataloging the open source holdings of the Community as a whole;

- establishing a comprehensive requirements system for the Community; and

- establishing interconnectivity so open source information can be shared throughout the Community.

Another responsibility of the Open Source Coordinator is to interact with the managers of the other collection disciplines to ensure that they are not collecting against requirements that can be satisfied through open source materials.

In my view -- and this is a view shared by many throughout the Community -- open sources should be the Community's first step in a range of choices to meet our overall information needs. Compared to information collected from satellite and other reconnaissance and surveillance means, open sources are relatively inexpensive to acquire. It would be both bad acquisition management and information management to waste a costly intelligence asset collecting information that can be acquired through an open source.

Although I believe that open source should be the Community's first step in attempting to satisfy our information needs, I want to emphasize that it will likely never replace the other intelligence collection disciplines. But I do strongly believe that better and complementary management of open source assets will, in turn, lead to more

efficient and focused use of those other collection disciplines.

Without question, the biggest challenge the Intelligence Community faces with respect to open source is processing the vast amount of data available.

Intelligence organizations have significant expertise processing and filtering large quantities of data, putting it on mass storage, manipulating it (including translation or gisting), and devising systems to have data available to analysts on an on-call basis.

In fact, US intelligence operates what is probably the largest information processing environment in the world. Consider this: Just one intelligence collection system alone can generate a million inputs per half hour; filters throw away all but 6500 inputs; only 1,000 inputs meet forwarding criteria; 10 inputs are normally selected by analysts and only one report is produced. These are routine statistics for a number of intelligence collection and analysis systems which collect technical intelligence.

In the open source arena, FBIS monitors over 3,500 publications in 55 foreign languages. And each day it collects a half a million words from its field offices around the world and another half a million words from independent contractors in the US -- that's equivalent to processing several copies of War and Peace every day.

These two examples show the magnitude of the classified data and translation problems facing an already data-rich Intelligence Community. The open source challenge can theoretically present ever more daunting levels of data and translation requirements, reaffirming that information management will be the single most important problem for the Intelligence Community to address for the future.

In this equation, one of the dilemmas posed deals with how we will spread our already overtaxed Intelligence Community information management resources in the context of both people and systems across both the classified and open source collection and analysis areas. This conference can hopefully provide some pointers to solutions in this area, mindful of the fact that we are in a period of intelligence budget austerity and Community downsizing.

Of course, the key issue for intelligence analysts is not simply the quantity of open source data that is collected, but also its quality -- that is, its intelligence value.

Open source information is produced or published largely according to the needs of the private sector, without regard to the uses to which that information will be put by the intelligence customers.

At the present time, open source materials that are collected for intelligence are often not made available to analysts in a way that is useful to them. And there is only limited ability to search large open source holdings in a timely manner.

A substantial amount of open source information is reported in foreign languages and require translation. For example, in FY '92, FBIS translated 200 million words-- so the translation issue is another dimension to the information management challenge.

Much of the Intelligence Community's current open source architecture was developed in an age when information processing and communication were in their infancy. As we look to the future, we will have to develop more creative approaches to manage the vast amount of data being produced.

Based on the recommendation of the DCI Task Force, the Community-wide Open Source Steering Council has developed a Strategic Plan that presents a vision of the Intelligence Community's goals for open source collection ten years from now.

The plan establishes the goal of creating an integrated Community open source architecture. The new architecture must provide, among other things:

- flexible collection,
- networked access to external data bases,
- immediate user and customer feedback, and
- automated, profiled delivery of collected open source information based on user requirements.

We expect the centerpiece of the architecture will be an Open Source Information Exchange -- comprising a central switch and digital communications networks which interconnect all user organizations within the Community.

The inputs to the Open Source Information Exchange will be government sources, such as the Library of Congress; the FBIS electronic dissemination system; and the vast array of commercial sources, such as NEXIS.

The system will distribute open source data through "functional support centers" that are being developed and

funded by the Community. These functional support centers will serve as focal points of expertise on critical intelligence topics -- such as science and technology; and political, military, and economic intelligence.

The strategic plan begins with a vision for timely access to open source data, and identifies funding, management, and architecture considerations. It also establishes strategies and outyear timetables for Open Source Coordinator pursuit of a more integrated open source design and implementation effort to support the widest variety of Community analysts and policymakers. Finally, it deals with specific requirements, collection, processing, exploitation, dissemination-related goals and objectives in specific detail. The plan is ultimately a roadmap for Intelligence Community-wide transition from the current way of doing business to future mode of operation. Its principal features bring us:

- from partial connectivity to full interconnectivity;
- from reliance on hard copy to digital electronic data;
- from dependence on physical information centers to the establishment of electronically-connected, virtual information exchanges;
- from querying one data source at a time to querying many at once;
- from the use of multiple standards and tools to common standards and tools;
- from development of sophisticated data bases on only a few topics to the creation of a wide range of data bases on many subjects;
- from limited access by users to sources of data to broad, flexible access;
- from emphasis on the collection and dissemination of data to a balance between that and making data readily accessible and tailored to user and customer needs;
- from a limited ability to display data to the ultimate application of fast developing multimedia capabilities.

Many important questions must be considered along the way. Open source is being considered by the DCI as the equivalent of a dominant intelligence discipline. In the restructuring of national intelligence, we have structural agency czars who are focused on and accountable for the

management of national disciplines; for example, NSA for SIGINT, CIO for Imagery and the CIA/DO for HUMINT. Open source is far more decentralized in its current and even its envisional approach and management. Will the time come when we need more focus and accountability in open source management, mandating the establishment of a structural czar for this critical area?

How will we change the mindset of those people in the Community who do not yet think of open source as a bona fide collection discipline on a par with SIGINT or IMINT, or HUMINT?

There is no question that open source -- in comparison with other collection systems -- has the potential to provide a lower cost, lower risk supplement to intelligence collection and analysis. But access to open source data still costs money. So another question the Community will have to consider is, "just how much money will be available to spend on managing open source data in an era of potentially dramatically declining resources?"

An expanded use of open source material raises legal questions -- especially concerning licensing agreements and copyright protection. Lawyers and managers in the Intelligence Community are working diligently to ensure that our use of copyrighted information strikes the appropriate balance between the government's legitimate need for access to open source material with the copyright owner's rights and privileges.

For instance, the Intelligence Community, like any other commercial user, buys access to a number of commercial data bases. When one component of the Community is licensed to use a data base, can it disseminate that data, or provide access to other users in the Community? If the answer is not clear, we work with our commercial vendors to revise our basic agreement to ensure that our use of the material is consistent with our agreements with the vendor.

I would like to conclude by reminding you that the great strength of American intelligence, which is unique in the world, is its ability to responsibly manage a global intelligence system, continuously moving bits of data from diverse sources to a broad and demanding customer-set. That new customer-set has even more highly distributed information requirements for the future.

Fundamentally, it is the hundreds of messages and other intelligence products that we electronically disseminate hourly which constitute the bread and butter of the intelligence business. Our implicit requirement is to manage a virtual intelligence system which adapts its multimedia

products to the demanding users in a changing and unsettled world environment. Open source material fits legitimately and prominently into the equation of modern intelligence sources and methods, but presents special challenges and dilemmas for us to resolve.

Well, that's a lot to think about -- and I can assure you that the Intelligence Community is well aware of the opportunities -- and the challenges -- associated with open source in our world today and of tomorrow. And we plan to draw on the expertise of the private sector, other government agencies, and the academic community to meet these challenges in the years ahead.

This conference is just one way we can help to make all this happen. Thank you for your interest in this important topic; we look forward to continuing to work with you in the future as we attempt to address these and other problems associated with the provision of focused intelligence and information support to our customers in a changed and changing world. Thank you for the opportunity to be with you today, and I wish you well for the remainder of this timely and important conference.

FIRST INTERNATIONAL SYMPOSIUM: NATIONAL SECURITY & NATIONAL COMPETITIVENESS: OPEN SOURCE SOLUTIONS Proceedings, 1992 Volume II - Link Page

Previous	OSS '92 Dr. G. A. Keyworth II, former Science Advisor to the President of the United States, Government as a Customer in the Digital Age,
Next	OSS '92 Dennis Clift, Defense Intelligence Agency, Military Open Source Requirements, Capabilities, and Contracting Directions,

Return to Electronic Index Page