View from the Private Sector

OSINT: The State of the Art, the Artless State[1]

Mark M. Lowenthal

> **If the IC were to understand OSINT better, it might get the emphasis it deserves.**

Mark M. Lowenthal is the Director of the Open Source Intelligence Program at SRA International. He is a former Staff Director of the House Permanent Select Committee on Intelligence.

This article is unclassified in its entirety.

One of the unique burdens of being a practitioner of and advocate for open-source intelligence (OSINT) is the constant need to define what OSINT is and why it matters. Even though these are not controversial issues, it is useful to begin on common ground.

Seen from the perspective of the Intelligence Community (IC), OSINT is usually regarded as "everything else," after the classified sources of IMINT, SIGINT, HUMINT and MASINT are accounted for. Each of these four collection disciplines differs from the others in what it collects and how it collects it. Because a significant portion of OSINT is unique, extracted from an array of textual sources, it has traditionally been considered a separate category of intelligence information.

In fact, however, OSINT is the most pervasive of the INTs, rather than a separate category. It occupies its own niche as well as some part of each of the other INTs. Beyond the textual sources of OSINT, the only aspect that differentiates it from other collection disciplines is the fact that it is not clandestine in nature. Many aspects of OSINT are quite the *same* as the other collection disciplines, because open-source intelligence is a facet of each of them, not an entirely separate discipline. It is more useful and more accurate to view the disciplines as the diagram indicates.

A New Way To Visualize The Collection Disciplines

CLANDESTINE	*OPEN*
IM	INT
SIG	INT
HUM	INT
MAS	INT
	OSINT

If the IC were to understand OSINT better, it might get the emphasis it deserves. Seeing OSINT as a facet of each of the other collection disciplines underscores the importance of open-source information to the all-source process. OSINT matters because it can be a cost effective, significant source of intelligence. OSINT should be seen as a daily gift, a leg up on each day's collection tasking. Instead, OSINT is seen as something less than a consolation prize—something done grudgingly, haltingly, and poorly in the IC.

The State of the Art

Open-source intelligence today is a much wider field than a few years ago. Most OSINT practitioners still start (and, too often, stop) with unique textual sources. These fall

[1] With an appreciative nod to Dr. Amrom Katz, whose article "Verification and SALT: The State of the Art and the Art of the State" (Heritage Foundation, 1979) inspired this title.

into two major and two minor categories.

The most important textual source, in my view, is commercial media vendors, two of whom predominate: Dow Jones Interactive and Lexis-Nexis. Both offer access to multiple worldwide sources of information via a website. These sources include newspapers, magazines, and some broadcast transcripts—all in English translations that are accurate and reliable. The websites allow specific searches by words and combinations of words, and by specific periods of time. It is possible to create permanent files of search terms so that they do not have to be reconstructed each day.

The second major OSINT text category is the worldwide web itself. Here the OSINT practitioner faces several problems:

- The first challenge is sheer volume: what were once thousands of pages of information have become billions. This may be the ultimate "needle in a haystack" problem. It is compounded by the fact that no two search engines will identify the exact same sites. A better approach is to use meta-search engines, i.e., search engines that are actually a combination of other search engines. For example, Copernic.com simultaneously searches Alta Vista, Excite, HotBot, Lycos, MSN Web Search, Yahoo, and several others. This type of search increases the chance of finding the best sites, although there still is no guarantee. Access to these meta-search engines is free via the Internet itself.[2]

- Another difficulty is source validation. The web is a vast electronic bulletin board where users can post anything they want. There is no restraint, no authorization, and no authentication. It is the First Amendment on a global scale, pushed to the point of anarchy. As with any other source, over time analysts learn which sources are reliable and which are not, but it is a pretty steep learning curve at the outset, with lots of chaff and little wheat.

"Open-source intelligence today is a much wider field than a few years ago. OSINT practitioners ...too often stop with textual sources."

The two minor textual sources of OSINT are hard copy material, primarily books, and so-called gray literature. Yes, books still matter in the "e-age." Using books, of course, usually requires visiting a library, and analysts are increasingly reluctant to do this. "If I can't get it on the web, it's not worth getting," seems to be the attitude. Gray literature is one of those phrases of art that means different things to different people. I think that some OSINT practitioners like to use the phrase "gray literature" because it gives a patina of "spookiness." It should not. The Fourth International Conference on Gray Literature, held in Washington in October 1999, defined this source as: "That which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers." This can be a useful range of sources, although it usually takes some search skills to hunt them down.

Those of us toiling in the commercial OSINT vineyards have seen some remarkable changes in open-source intelligence in the last few years. In additional to information from unclassified textural sources, OSINT now pervades all of the collection disciplines.

- **Open-source IMINT:** The arrival and development of commercial imagery down to resolutions of less than one meter is still a breathtaking and somewhat scary development. Once the exclusive intelligence domain of a few technically advanced states, high quality imagery is now available to anyone who can afford to buy it.

- **Open-source HUMINT:** The open elicitation of information by tasked collectors can be an incredibly useful source of intelligence. We are *not* discussing breaking and entering, bribery, or any other activity that is illegal or even questionable. We are discussing collection that is no different from what foreign service officers or defense attaches do every day. Indeed, I would

[2] Some web searchers like to use "web crawlers," in effect, search devices that look for specific words or phrases automatically. I have found these to be generally unsatisfactory. If the parameters are defined too narrowly, the returns are thin. Broaden the parameters and one faces inundation. Modulation tends to be difficult and imprecise.

consider much of their reporting to be open source, despite its near automatic classification as SECRET, since their interlocutors know that what they say will be reported back to Washington.

- **Open-source MASINT:** Simply put, if you can purchase commercial imagery, you can purchase "imaging" that uses part of the spectrum other than the visual range. This may not be available yet across the full multi- and hyper-spectral range, but there is little reason to doubt that it eventually will be.

- **Open-source SIGINT:** Information can be gleaned from communications without resort to wiretaps or other means of intercepting signals. As just one example, on the worldwide web one can use programs to analyze traffic on a site and monitor changes in content from day to day.

Need for a Methodology

As is evident from the foregoing, there is a great deal of OSINT out there. This underscores the importance of *process*. As with any collection discipline, OSINT needs a methodology for dealing with it. My own preference is the approach used by the IC itself: requirements, collection, processing and exploitation, and analysis. The first two steps will come as no surprise. You need to know what you are collecting and why (requirements) and you need a plan for meeting this requirement (collection). It is important to note here that not

"I believe that one of the reasons why OSINT has failed to reach its potential in the IC is because responsible officials forgot that technology was the means and content the end."

everyone can collect OSINT. Even people who surf the Internet may not be skilled at open-source collection. It takes a specific set of skills and a specific analytic outlook.

OSINT does not arrive full-blown and usable from *any* of its sources. It requires some level of processing and exploitation. For open-source IMINT, SIGINT, HUMINT and MASINT, the need for processing and exploitation is obvious. For OSINT, although less obvious, even the traditional textual sources require vetting, validation, interpretation, sifting, and weighing. The degree to which the exploitation process can be automated is still difficult to determine. Numerous firms offer technologies that perform text mining, text summarization, or link analysis. In the proper context, these tools can be extremely helpful to the analyst slogging through mountains of open-source information. They cannot simply be thrown at the problem, however. There has to be a plan—the right tools for the right job. It is important to avoid the American love affair with technology for its own sake. Clients' intelligence needs (whether government or commercial) cannot be met by merely placing a search engine or text miner before them. The OSINT requirement will be met by *content*. Technology is a means to that end, not the end itself. Once the information is sifted and organized, analysis of the material becomes possible.

OSINT and Commercial Clients

Providing OSINT to large commercial firms differs considerably from serving IC clients. It remains a nascent industry for several reasons:

The concept of intelligence is inherently less familiar to industry than it is to the IC or to the policy agencies that are served by the IC. One must spend more time explaining what intelligence is, and what it is not.

It is difficult to determine the right point of entry into large firms to make OSINT sales. The president or CEO is too high. Directors of research see OSINT as little more than a threat to their own activities. Other levels and individuals need to be probed. It is a more random effort than dealing with the IC.

Although commercial users of OSINT want "actionable" intelligence, just like US officials, this desire may not match their actual business practices. Large firms do not make key decisions with the same frequency as government officials. Rather, they have sales campaigns that have beginnings, mid-year and end-of-year reviews, and renewed starts. The provision of intelligence on a regular basis

does not fit into this less frequent decisionmaking structure.

Commercial clients appear to be less comfortable dealing with the uncertainty that remains *after* intelligence has been provided. This may be a result of a general unfamiliarity with intelligence and overly high expectations as to what it can do. Again, to avoid or to overcome this problem, it takes tutoring on setting reasonable expectations.

Although there are many OSINT providers who espouse the cause of competitive intelligence, few of these firms or individuals have any real sense of what constitutes intelligence or an intelligence process. They seem more concerned with constant self-definition as a professional group than with the application of proven methods to their own activities.

That said, however, there are some interesting and challenging opportunities in the commercial OSINT field. Indeed, many of us find these more attractive than working with the IC, because they are more challenging and, to be frank, more lucrative than government contracts. Some of us also believe that our efforts to apply the methodologies we have learned in the IC to commercial clients will eventually pay off in enduring relationships. The commercial engagements also provide opportunities to experiment with new methodologies and new techniques and tools that can then be added to offerings to the IC.

"To be very clear, I do not view OSINT as a panacea for all intelligence requirements."

The Artless State

Given the rich, evolving OSINT world, why does the IC continue to lag in the use of open-source information? IC defenders of open source will point out, correctly, that the IC has always used OSINT. This is undeniably true. But it is also true that the IC's use of OSINT remains fitful and sometimes grudging, and in no way fully exploits the available opportunities.

I believe part of the reason is institutional. If all of the intelligence we needed were available openly, then we would not have an IC *per se*. We would have something else—a research center of some sort. The need to collect *some* intelligence clandestinely, however, leads to an ethos in the IC where clandestine intelligence is consistently valued more highly than open-source information. As currently practiced by the IC, open source is, at best, an add-on to clandestine collection, not a partner.

To be very clear, I do not view OSINT as a panacea for all intelligence requirements. Nor do I believe that the increased availability of OSINT ends the need for various covert sources and methods. There will always be intelligence requirements that can only be met clandestinely. At the same time, we now have more intelligence requirements that can be met openly—if the IC is willing to go in that direction.

The IC's exploitation of OSINT also has been handicapped by the problem of technology seduction noted above. The Community has prospered, in part, because of various technologies it has created—primarily in the collection area. I believe that one of the reasons why OSINT has failed to reach its potential in the IC is because responsible officials forgot that technology was the *means* and content the *end*. They pursued technological solutions, but many failed to deliver coherent, useful content to analysts, which should have been the goal.

Broader institutional problems within the IC also figure in how OSINT is handled. I will illustrate by anecdote. I had a contract with an all-source agency to provide OSINT about Bosnia. In a moment of self-abnegation, I pointed out to the agency that they already had Lexis-Nexis service and could obtain on their own some of the OSINT I planned to provide. I will lapse into dialogue to convey the exchange:

All-Source Agency: "Oh, we never use Lexis-Nexis for Bosnia."

Lowenthal: "Why not?"

All-Source Agency: "Because that would reveal that we are interested in Bosnia as a requirement."

Lowenthal: "We have some 9,000 troops in Bosnia. They rotate in and out to great publicity. Don't you think people know we are interested in Bosnia?"

Again, the cult of secrecy gets in the way of what would appear to be a sensible use of OSINT. I have to believe that a community that creates cutouts and false fronts could figure out how to use OSINT without revealing intelligence requirements.

The use of commercial imagery is another interesting illustration of the IC's shortcomings in dealing with new OSINT opportunities. The National Imagery and Mapping Agency (NIMA) expects to be the sole purchaser of commercial imagery for the government, or at least for the IC. NIMA's goal is to prevent duplicative purchases, which is laudable. NIMA, however, has *no* interest in promoting greater use of commercial imagery, which it views as competition. It would be far better to give various would-be commercial imagery users their own budget, with NIMA acting as a clearinghouse, developing and maintaining a master list of what has been purchased and is on file to avoid duplication.

The way the IC handles OSINT is also bizarre. For the other INTs, the IC has designated collectors, processors, and exploiters. For OSINT, beyond the Foreign Broadcast Information Service (FBIS), it has none. All-source analysts are expected to act as their own OSINT collectors ("Go surf!"), as well as their own processors and exploiters. To comprehend the ludicrousness and self-defeating aspect of this, imagine an analyst working on Iraq being told: "We have a lot of images from Israel to Pakistan and yours are in there somewhere. And we also have

"The fact that no one in the IC is responsible for OSINT, and no one seeks to be, speaks volumes."

5,000 SIGINT hits from Iran and Iraq. Good luck." That is not far different from how the IC handles OSINT. It is self-defeating and says much about how the Community views OSINT.

The clearest indication of the sad state of OSINT in the IC, however, is the fact that no one is responsible for it. If I have an IMINT problem, I go to NIMA; if I have a SIGINT problem, I go to the National Security Agency; if I have a HUMINT problem I go to the CIA or the Defense HUMINT Service; even poor MASINT has the Central MASINT Office. To whom do I go if I have an OSINT problem? No one. In Washington, bureaucrats have a nose for which programs will enhance their power and position and which programs are drags. The fact that no one in the IC is responsible for OSINT and no one *seeks* to be speaks volumes. OSINT is not worth fighting for or acquiring within the IC.

The Role of FBIS

Currently, FBIS is the centerpiece of OSINT in the IC. FBIS has been struggling and declining for years under repeated budget cuts, a sure sign of bureaucratic weakness. Strong programs prosper. FBIS has been providing open-source media coverage for as long as we have had an IC; however, FBIS's offerings are now outmoded compared to commercial vendors. These vendors are more expensive than FBIS, but the information they provide is more extensive and more timely. In my own open-source practice, I use FBIS occasionally as a supplement, but it is not my information supplier of first choice. I get better coverage from more sources more quickly by going elsewhere.

What Is To Be Done?

First, someone in the IC has to be responsible for OSINT. FBIS is too slender a reed, for the reasons already noted. I do not recommend reviving the Community Open Source Program Office (COSPO), which the DCI properly disbanded in 1998. COSPO, in my view, pursued the technology chimera and forgot about content delivery to analysts. Within the current IC structure, I would designate an OSINT Manager. The OSINT Manager would be responsible for creating a program designed to deliver usable OSINT—i.e., OSINT that has been collected, processed, batched, and sorted—to all-source analysts. I would place the OSINT Manager under the Associate Director of Central Intelligence for Collection, rather than in a specific agency. OSINT should be a service of common concern for every all-source analyst, not something controlled by one agency.

The OSINT Manager should oversee two core activities. One should be to seek the best commercially available technology and commercial vendors to "make the OSINT happen." The use of commercial off-the-shelf technology is important. There is no sense in

reinventing the wheel, another IC failure in open source. The second—and even more important—core activity should be to determine the OSINT requirements for each analytic office. I would start small, handling offices with discrete responsibilities first. I might recommend beginning with countries of relatively low day-to-day importance to the United States, where sudden crises can catch Washington off guard and where OSINT can handle a very large part of the intelligence requirements—most of which are not being met at all currently.

Happily, the dissemination problem is already solved. Intelink provides an up and running electronic means to broadcast usable OSINT to whomever might need it. It should be easy to create an OSINT portal for Intelink, where open-source information can be posted by issue—either regional or topical per the usual IC taxonomy.

Second, to aid in fixing the IC's OSINT problem I would liberate commercial imagery. Each agency should have its own budget to buy the imagery it needs, after checking with NIMA to make sure that a particular image has not already been purchased. Third, the IC's goal should be to collect OSINT *first*, not *last*, validating it whenever possible with classified sources. OSINT would become an enabler, not a niche filler.

Will all of this take money? Of course it will—and I am well aware of the constraints posed by the IC budget. Ironically, OSINT is a potential money *saver.*

- A well-integrated, content-focused OSINT program would be the most *efficient* collection management system we have had, because the IC would know what is available openly and could concentrate its classified collection in those areas where only it will suffice.
- All-source analysts would have more time to analyze, avoiding the frustration and valuable time it takes to do their own OSINT collection, processing, and exploitation.
- The IC would have better intelligence on many of the lower priority issues, where little of what we need to know requires clandestine collection and where we currently devote few collection resources at any rate. These are often the issues that become the "surge" problems, when a Burundi or a Somalia explodes and we scramble for even the most basic information.

Utopia? Not really. It would simply be a smarter use of open-source information than we have at present. It is less an issue of money, know-how, or technology, than it is one of will. OSINT continues to grow and prosper. The IC has yet to take full advantage of it.