Crew chief with 36th Aircraft Maintenance Unit, Osan Air Base, South Korea, checks computer data during Red Flag-Alaska 14-2, ensuring F-16 Fighting Falcon readiness (U.S. Air Force/Peter Reft)

Defense Intelligence Analysis in the Age of Big Data

By Paul B. Symon and Arzan Tarapore

Over the past decade, the U.S. and Australian intelligence communities have evolved rapidly to perform new missions. They have developed new capabilities and adapted their business processes, especially in support of joint and complex military operations. But in the coming decade,

their greatest challenge will be to develop new capabilities to manage and exploit big data.[1] We use the term *big data* to mean the exponentially increasing amount of digital information being created by new information technologies (IT)—such as mobile Internet, cloud storage, social network-

ing, and the "Internet of things"—and the advanced analytics used to process that data. Big data yields not simply a quantitative increase in information, but a qualitative change in how we create new knowledge and understand the world. These data-related information technologies have already begun to revolutionize commerce and science, transforming the economy and acting as enablers for other game-changing technology trends, from next-generation genomics to energy exploration.[2] In defense intelligence communities,

Major General Paul B. Symon, Australian Army, served as Director of the Defence Intelligence Organisation from 2011 to 2014. Arzan Tarapore is a Doctoral Student in the War Studies Department at King's College London. The authors thank Josh Kerbel and Peter Mattis for comments on an earlier version of this article.

some of these technologies have been adopted for tasks, including technical collection and operational intelligence fusion—but big data's impact on all-source intelligence analysis has scarcely been examined.

This article offers a view on how these disruptive information technologies could transform defense intelligence analysis and the functions of the all-source enterprise. It is not a comprehensive study on trends in technology or in the intelligence profession, nor is it a deterministic scenario of a high-tech future. Rather, here we seek to identify some opportunities and risks of the disruptive technologies at hand. First, we sketch a background of the most important IT trends that are shaping today's economy and society. Second, we outline how big data could transform intelligence analysis; it has the potential to unlock enormous productivity gains and effectiveness by automating some currently labor-intensive tasks, enabling new forms of analysis and creating new forms of presentation. Third, we argue big data cannot do it all; its utility in making sense of complex systems and addressing knowledge gaps is limited. Finally, we outline how big data could transform the wider assessment agency enterprise. We argue that the explosion in data supply and demand will incentivize assessment agencies to reposition their roles more toward service-delivery functions and to rebalance their workforces.

None of this is inevitable. In both analytic operations and enterprise management, much of how the scenario actually unfolds will be determined by the vision and agility of our leadership, our partners, and our adversaries. Defense and intelligence community (IC) leaders must play an active but balanced role, exploiting big data's potential, but understanding its limitations.

Today's Tech Trends

The big data phenomenon presents defense intelligence with a range of opportunities, from off-the-shelf tools to complex business-process reforms. Some tools can be absorbed wholesale by the IC; for example, social network-ing tools such as Wikis and Chat are already being used to facilitate better collaboration between analysts. Beyond simple software acquisitions, however, disruptive information technologies have birthed a number of trends in how data are collected, moved, stored, and organized. Four of the most salient prevailing concepts, which are already transforming the economy and society, could reshape all-source intelligence.

Everything Is Social, Mobile, and Local. Much of the explosion of big data has been driven by the fact that information is increasingly social (generated and transmitted by many users, rather than a few big producers), mobile (collected by sensors on ubiquitous Internet-connected mobile devices), and local (geospatially tagged). These trends have irreversibly transformed IT; mobile devices in particular have become the primary means of connecting to the Internet and have thus become the primary market for much IT innovation. This has already created new opportunities not only for collection, but also for intelligence processing, exploitation, and dissemination (PED), and analysis.

Data Are Useless Without Data Science. The exponential creation of digital data holds enormous potential for creating insight and knowledge through PED and data analytics. The burgeoning field of data science—at the intersection of statistics, computer science, and other related fields—is increasingly being used by the private sector to realize the commercial potential of big data, often for prosaic tasks such as tracking a person's consumption patterns to better target advertising campaigns. The IC's routine work of collection, PED, and analysis is still largely organized on the Cold War model of seeking out sparse and secret information. Now, however, it must cope with the inverse challenge (and exploit the opportunities) of managing and analyzing massive quantities of data and, in the process, compete with the lucrative private sector to attract the highly specialized skills of data scientists.[3]

IT Solutions Are Customized and Intuitive. The accelerating pace of innovation and the need to best harness big data are both enabling and driving the creation of IT solutions that are customized and intuitive for the user. Gone are the days of hefty user manuals or obscure text-based user interfaces. Specific applications perform specific functions. Even major platforms such as Palantir are delivered with bespoke service support, both in tailoring the product to customer requirements and in providing ongoing software development support. Complex data-driven analysis demands a menu of apps or even dedicated software developers integrated into analyst teams—as they already are in some parts of the IC.

The Internet Is Everywhere. The rate of increase in big data will only grow as more devices join the Internet. These devices not only provide an interface for users, but are also creating a growing "Internet of things"—everything from household appliances to industrial robots—that generate and use more data, in turn creating more potential knowledge and vulnerabilities. At the same time, emerging technologies (such as free-space optical communications, which use lasers to transmit data through the atmosphere) are allowing users to bring the Internet into austere communications environments in order to enable the wider military use of Internet-connected IT and greater resilience to network failures.

These technology trends have been driven by the commercial and scientific sectors, but they also have powerful implications for the IC; they are rapidly challenging long-held conceptions of intelligence collection targets, business processes, required IT tools, and workforce skill sets. But the IC's capacity to adopt these technologies remains inadequate; fully exploiting these trends would require a deep revision of innovation policy and IT-acquisition business models. To adequately exploit these opportunities, the IC would need to incorporate a "technology push" acquisition model alongside the customary "demand pull" model. In today's IT environment of faster innovation and more disruptive and unpredictable technologies, where government lacks the speed or vision to lead innovation, the IC's best option may be to monitor

and leverage incipient innovation instead of attempting to drive it. Rather than dictating requirements to firms through a byzantine acquisitions process (as in most defense procurement programs), the IC's greatest potential for IT adoption may lie in injecting its "use cases" (and resources) in the start-up or development phases of future technologies. And in a data-intensive information environment, assessment agency leaders would need to recognize that adaptive IT is integral to analytic operations and no longer an ancillary support function toiling in the basement. The analysis mission-owner should therefore be responsible for shaping the agency's IT architecture as never before.

Even if imperfectly realized, today's technology trends hold enormous potential to transform all-source intelligence.

Transforming Analysis

Across intelligence problems, big data's greatest promise is its potential to integrate and organize information. New technologies for collecting, moving, storing, and organizing data could give all-source analysts access to vastly more information with more automation and productivity, thereby allowing them to concentrate their finite cognitive capacity on the hardest, highest-priority problems. But rather than simply bolting new technologies onto current processes, assessment agencies now have an opportunity to incorporate new technological trends in ways that fundamentally reshape how data are used for all-source analysis. The new technologies could be usefully applied to a range of defense intelligence problems, including social network analysis, weapons systems modeling, trend analysis for tactical military intelligence or nontraditional warning problems, and nascent analytic constructs such as "object-based production" and "activity-based intelligence."[4] Thus, they not only improve our capacity to execute existing intelligence missions, but they also create entirely new data-intensive types of analysis.

More Information with Less Effort. Big data and data analytics rely heavily on automation. Once the architecture

and algorithms are set, the data could be managed—collected, moved, stored, and organized—with relatively little additional effort. Applied to all-source intelligence, the exponential increase in data and analytics would render manual information retrieval impractical and unnecessary; the heavy lifting of data management could be largely automated. Already-existing tools can create an automatic and persistent push of data to analysts, obviating the labor-intensive requirement to manually pull data from various sources. That push of data could be more processed and valuable—for example, collated across different sources or formats—before it even reaches the analyst.

Automated data collation and analytics would both save analyst effort and enable powerful new capabilities. Data analytics could, with varying levels of human supervision, characterize data into meaningful clusters or categories, categorize and file new data into existing clusters, and detect outliers or new data that do not fit into existing clusters.[5] For all-source analysis, new methods such as object-based production could enable seamless integration of data from multiple sources and in multiple formats, thereby building comprehensive libraries of data on given targets. Analysts could use that mass of data and associated analytics to more quickly identify intelligence gaps, unexpected correlations and associations, or anomalies or irregular behavior. This range of capabilities could be profitably used, for example, for everything from finding patterns or anomalies in a terrorist target's pattern of life, to tracking military targets automatically in wide-area surveillance, to tipping and cueing for humanitarian assistance and disaster recovery support. In such cases, human intervention—especially expert analysis of the target—is still critical, but big data could empower those analysts to know more and to know it more quickly and with less effort.

Big data technologies allow intelligence to move quickly, be stored indefinitely, and yield more valuable insights over time. Much of the newly collected data would arrive at or near real-time, compressing the latency of

collection, PED, and analysis, and cueing further collection. Vast quantities of data—unprocessed and unseen by any analyst—would be stored, available to be mined later in the context of future data or requirements or to discover or recognize associations or trends. Machine learning would allow this entire process to improve with time. The accumulation of data and the refinement of algorithms would allow for dynamic and progressively more accurate models or more robust and adaptive normalcy patterns, and would enable the detection of finer or more meaningful anomalies accordingly.

There are significant challenges to fielding these new capabilities. Some of these challenges are technical—for example, optimizing ways to ingest and collate data from different sources and in different formats, especially unstructured data from text and media. The thorniest challenges, however, are associated with policy settings and governance frameworks. For example, intelligence agencies will need to set standards for the vetting and quality assurance of data they source from interagency or other partners; establish security and legal compliance protocols for sharing data across organizations; establish robust security measures to protect data from spoofing, cyber exploitation, or insider leaks; and standardize the tagging and coding of data for use in analytics. Once mission-owners set these frameworks to govern the effective and secure use of big data, all-source analysis should yield unprecedented gains in productivity and capability.

Presentation Is Everything. Once collated, managed, and applied to gain new insights, data must be presented effectively to the customer. Here, too, big data carry risks and opportunities. Customers will never lose the temptation to acquire and interpret their own data, and big data, plentiful and apparently authoritative, will exacerbate that problem. The IC faces the risk that these quantities and varieties of data will create the appearance of veracity—and customers' easy access to raw data streams or intelligence reporting could become even more hazardous. In an environment where data are

Airman checks diagnostic information after applying three different upgrades that give pilots more situational awareness data in user-friendly formats (U.S. Air Force/Alexander Guerrero)

ubiquitous, customers will expect immediate and authoritative answers and will sideline IC producers that cannot quickly deliver user-friendly products.

Fortunately, big data and data analytics also present opportunities to create compelling and effective outputs for the customer. Data-intensive solutions to intelligence problems demand appropriate forms of presentation; just as in science and commerce, these solutions would be best presented as graphics or visuals, not text-heavy assessments. Assessment agencies could profitably use one or a few main data-agnostic platforms (such as Google Earth), connected to relevant intelligence databases and easily overlaid with various customized data layers, to electronically deliver finished intelligence to the customer. With the concomitant improvement in IT, these outputs could be easily pushed to the customer, just as data are pushed to the analyst. Presented in multimedia, they could incorporate

multi–collection platform reporting and data streams and use "recommendation engines" of the type used by Amazon and Netflix to suggest other relevant outputs tailored to the customer's requirements.

The most effective finished intelligence outputs, exploiting the full potential of data analytics, would incorporate the following features. First, they would use a visualization platform, and for strategic analysis, the most common platforms would most likely be geospatial. Much digital data are already geospatially tagged, and geospatial presentation often yields powerful insights that are not otherwise apparent. Second, they would be dynamic—using automated feeds, the product would be constantly updated with data collated in real time. Outputs would offer more than just a recent snapshot of intelligence, as the IC typically provides now with written assessments, and they would render obsolete terms such as "Latest Date of Intelligence" or

"Information Cut-Off Date." Third, they would be interactive; the customer could interrogate the product, using hyperlinks or some other intuitive interface, to pursue additional layers of data.

These attributes of data-intensive presentation are clearly better suited to some outputs, and some customers, than others. Already, strategic assessments for national policymakers can profit from visual and interactive outputs—even the President's Daily Brief, the pinnacle of national-level intelligence, has been delivered on an iPad. With time, big data and data analytics could transform all phases of analytic operations, culminating with quicker, more accurate, and more tailored intelligence for customers.

Limits to Transformation

The promises of big data are tantalizing, but they are limited. The greatest impact will be felt in the analysis of who, what, where, and when ques-

Intelligence analyst gives commander of 21st Theater Sustainment Command terrain brief of Hohenfels Training Area on enemy activity (U.S. Army/Henry Chan)

tions, using single– or multi–collection platform structured data to address discrete, bounded questions. It plays a smaller role in analysis of why or how questions, which are salient not only for strategic intelligence supporting the policymaker, but also for every level down to tactical intelligence supporting subunit commanders.

Analysis Needs More Than Data.
Data-intensive forms of analysis promise new efficiencies and insights, but at its heart, all-source analysis needs more than just data. First and foremost, analysis needs expert leadership. Faced with the allure of compelling data, the IC faces a risk that available data will drive the analytic agenda rather than the other way around. The sheer availability of certain types of data could skew the analytic enterprise to prioritize its efforts or distort its assessments by placing undue importance on the most data-intensive activities or by emphasizing the most visible and trackable targets or issues. Instead, expert leadership must still determine which data are collected and in the service of which analytic priorities; these tasks demand judgment and knowledge of customer requirements. The analysis mission-owners must be careful to redouble their emphasis on directing the intelligence cycle

and to ensure the enterprise is serving customer requirements—asking the right questions and directing collection and analysis accordingly—rather than being slaves to the data.

Second, analysis needs expert analysts. Data-intensive fusion, PED, and analysis are better suited to some types of intelligence problems than others, but they always require expert analysts to make sense of outputs. Data-intensive analysis can more profitably be applied against "puzzles," with bounded, empirically discoverable answers, rather than "mysteries" that deal with a contingent, imponderable future.[6] Puzzles typically relate to discrete objects—places and things—whereas mysteries are tied to complex phenomena.[7] Mysteries or complex phenomena are the product of inscrutably complex human interactions and, like any complex system, are sensitive to countless variables and therefore inherently unpredictable. Defense intelligence must be postured to tackle both.

Even puzzles require expert analysts—to frame the puzzles in the first place, solve them, and then to make them relevant. Analysts need to verify collected data that may be flawed or spoofed by denial and deception, which requires expert analytic tradecraft. They then

need to provide the necessary context or value-added interpretation of the data analytics—the "so what?"—which requires not only subject matter expertise but also sensitivity to customer requirements.

Consider the conflicts that flared in Ukraine and Iraq in 2014. In both cases, irregular forces—Russian-backed separatists and Islamic State militants, respectively—made rapid advances against their adversaries, not only deploying effective military force but also documenting their campaigns in social media platforms such as Twitter and YouTube. Exploiting the content and metadata of these sources, fused with data from traditional intelligence, surveillance, and reconnaissance (ISR), could yield significant data about those forces' tactics, social networks, and geolocation at particular times. Those data-intensive streams would allow Western defense intelligence to build a high-fidelity picture of these forces' composition, materiel, and disposition. They could thus provide useful context and cueing for tactical intelligence support. But they would add little to the customers' understanding of the militants' intent—their operational plans and political agenda—or even some elements of their capability, such as their level of unit cohesion. Framing, solving, and interpreting these puzzles, even for tactical military intelligence problems, require analytic judgment, attuned to customer needs.

For mysteries, data may offer valuable piecemeal insights, but expert analysts need to do even more heavy lifting to translate those insights into meaningful assessments for customers. Expertise is critical for inferring a target commander's intent (as in the Ukraine and Iraq irregular warfare examples above) and even more so for assessments of complex phenomena, such as political unrest. For instance, a more perfect data-intensive coverage of the Arab Spring unrest could have provided better insights into the depth of popular opposition to Arab regimes or tactical warnings of intensifying protests, but simply a better coverage of social-networking or other data-intensive tools would not have prepared Western intelligence agencies to anticipate the

revolutions. Twitter feeds alone could not explain why revolutions swiftly consumed regimes in Tunisia and Egypt, or explain the difference in political trajectories in Libya, Bahrain, and Syria. An actionable intelligence response to Arab unrest would have required marrying that data-intensive coverage with subject matter expertise, comprehensive analyses of state stability, and a receptive and agile policy customer; big data without those factors would have provided tactical tipping of protests, not strategic warning of regime collapse or civil war. For complex problems, big data can provide a more granular picture of the target, quickly and with little effort, but the mystery can only be anticipated or managed (if at all) by the enterprise's expert leaders and analysts, working closely with the customer.

Addressing Knowledge Gaps. Some big data proponents argue that new storage and processing technologies should allow users to collect and manage virtually all relevant data about a given object. By examining the entire population of data rather than a sample (that is, where n = all), users could make direct observations rather than relying on inferences based on partial data. Induction and modeling would be unnecessary, replaced by the volume and fidelity of a virtually complete data set, manipulated by well-tested algorithms. In this view, better understanding only needs better data.

The quest for more data is all too familiar for the Intelligence Community. Built in the Cold War, when clandestine collection was key to uncovering scarce information, and reinforced in the past decade of ballooning technical ISR collection to support warfighters, the community has developed as a collection-centric system geared toward plugging intelligence gaps or arithmetically connecting the dots, and any missteps or intelligence failures are most commonly met with demands for more or better data.[8] For some problems, addressing intelligence gaps is vital, and big data will help—with both open source and intelligence collection.

Complex phenomena, on the other hand, are not so easily conquered by data. For these, assessment agencies need to address enduring *knowledge gaps.* Unlike intelligence gaps, knowledge gaps have no single, durable answer and may not be required to directly support specific decisions or actions. Rather, they are an ongoing requirement, a framework to guide collection and to improve decisionmakers' understanding as they seek to execute a plan. These gaps would only be satisfied—or, more likely, de-prioritized—when they are no longer essential for decision advantage. More data cannot close a knowledge gap. As a result, knowledge gaps involve an inescapable degree of uncertainty and limit analytic confidence. They remain extremely useful constructs to structure and prioritize intelligence collection and analysis, but they also highlight the limitations of big data's utility to strategic analysis.

Knowledge gaps may be comprised of multiple intelligence gaps, but critically, they also require analytic interpretation and judgment. For example, cataloging the signatures of China's new aircraft carrier, charting the performance of its aircraft and weapons systems, or tracking its position on a patrol all represent intelligence gaps with discoverable answers. But understanding how that vessel might be used by Beijing, in concert with other capabilities in a crisis or as part of a coercive strategy, would represent a complex knowledge gap comprised of many constituent intelligence gaps and unknowable future courses of action that are contingent, complex, and unpredictable. Data cannot reveal what does not yet exist, such as adversary decisionmaking in a crisis. For such knowledge gaps, collecting and collating all relevant data would not be sufficient; better data may provide richer evidence for interpretation and anticipation, but it would only be a supplement to subject matter expertise and rigorous tradecraft.

In defense intelligence, creating knowledge requires more analyst effort than closing intelligence gaps, but it is also more important, at least to strategic policy customers. Making sense of complex systems and phenomena—creating knowledge—is central to sound decisionmaking. Some big data optimists suggest that uncovering

A Modest Time Horizon

The pace of technological innovation is extremely high and increasing. Many of the consumer products and underlying technologies that have revolutionized the high-tech sector have gone from prototype to ubiquity in a few short years. The iPhone—the device that made mobile Internet routine—was launched only in 2007. By 2008, the number of mobile (WiFi) broadband users overtook the number of fixed (wired) users. The two giants of social networking, Facebook and Twitter, were both opened to public use in 2006; by 2011 Twitter was being credited with facilitating political organization in the Arab Spring, and by 2012 Facebook boasted more than a billion members.

These technology applications all had a widespread disruptive effect in less than 5 years; other technologies such as the "Internet of things" are yet to mature, and their impact can scarcely be predicted. Big data technologies present a complex set of challenges that the Intelligence Community (IC) will grapple with for years, but the extreme pace of technological change will continue. Within another 5 to 10 years, the high-tech ecosystem will probably be unrecognizable, and the IC will be faced with a radically different set of risks and opportunities. Thus we can only meaningfully project the impact of existing disruptive technologies with a maximum time horizon of about 5 years—a period that will be dominated by adoption of big data technologies. Anything beyond that is science fiction.

all relevant data for a problem (or achieving n = all) should allow users to draw reliable empirical correlations without needing to understand causality; indeed, in some fields, that may be sufficient. But in intelligence analysis, understanding causality is indispensible because customers seek to take action

to influence outcomes, and actionable intelligence support should accordingly highlight causality, enable the customer to understand their points of leverage, be alert to key decision points, and act effectively against threats or opportunities. Understanding causality in the context of customer requirements—in other words, creating and applying knowledge—is thus central to the IC mission.

Transforming the Enterprise

Simply passing the deluge of data on to customers would be counterproductive; even neatly presented fused data, absent expert assessment and advice, would only decrease the signal-to-noise ratio of useful, actionable intelligence. Big data are exacerbating that problem by sharply increasing both the *supply* of data available to the IC and the *demand* for it from senior customers. Caught in the middle, IC leaders will need to adapt not only to the transformation of analytic operations, but also to the functions and staffing of the enterprise.

From Production to Service Delivery. In an environment of ballooning data inputs and expected outputs, the IC cadre of all-source analysts will find it increasingly difficult to remain the original producers of all finished intelligence for their customers. Even with the anticipated productivity dividends, the enterprise in its current form will not be able to cope with the pace or scale of the big data challenge, for at least four reasons.

First, customer expectations are already growing and outstripping the IC capacity to adapt. As their decision cycles continue to be compressed, customers will demand immediate and data-rich answers rather than lengthy deliberations or vague and unverifiable "gut calls."

Second, in the face of these increasingly unforgiving expectations, the current production process—tasking collectors, collating and analyzing data, and producing finished intelligence reports—is too cumbersome and time-consuming. If the IC rigidly sticks to that process, dissatisfied customers will seek their information elsewhere.

Third, these dissatisfied customers will find data-intensive information support from a proliferating array of competing suppliers, from established and nontraditional media to commercial intelligence services, which can provide quicker and more user-friendly answers—at a tiny fraction of the IC enterprise's operating budget.

Fourth, the proportion of useful information that is classified, the unique province of the IC, is rapidly declining. Increasingly, decision advantage hinges on speedily integrating multiple streams of data rather than on a well-placed spy—and big data provide a wealth of open source or gray information that can more cheaply and automatically be deployed for intelligence solutions. Classified collection will remain indispensible, but IC leaders will be incentivized to more judiciously deploy those relatively expensive and risky means against their toughest hard targets.

With these clunky production processes, tough competitors, and less unique information, an unchanging IC enterprise will face an urgent threat of irrelevance. This threat sharpens already existing incentives for assessment agencies to reimagine their function, from the current industrial-age model of linear *finished intelligence production* to an information-age model of integrated and adaptive *assessment service delivery*. Even without the advent of big data, a growing body of literature on the state of the art of all-source analysis argues that intelligence agencies should cultivate a more intimate relationship with their customers—to better understand their requirements and more effectively deliver influential support—and to reconceptualize their role from sole producers to service providers.[9] Much of this literature points to the importance of timely and tailored on-call expertise (as distinct from discrete written products) as a key service for customers. The J2 briefing the commander or the analyst briefing the policymaker is an indispensable face-time moment for both the customer and the intelligence provider. The customers' abiding preference for agile and responsive in-person expertise will ensure such services remain a prized feature of assessment services.

Another key service the enterprise could deliver is access to a much wider network of expertise from across, and from outside of, the IC. In this view, assessment agencies would retain their core analysis and production mission, but to meet customers' demand with the best possible intelligence support, they would also leverage networks of other agencies, allied partners, commercial sources, and cleared outside experts. In a world awash in data, assessment agencies' prime advantage will lie in the privileged access to their customers; while they will not be able to internally produce all the answers, they should be able to tailor and fine-tune intelligence solutions sourced from intelligence collectors and from elsewhere. This service then amounts to enterprise management: using networks of experts and data sources and collaborative mechanisms including social-networking tools to quickly address priority knowledge gaps. Effective enterprise management hinges on robust integration with both those networks and with the customer.

Renewed Importance of Staff Functions. All-source analysts have traditionally been the core skill set of assessment agencies, and as we have argued, big data create powerful reasons to integrate data scientists and software engineers into analytic teams. Additionally, intelligence staff functions—a greatly enabled version of today's collection managers as distinct from all-source analysts—would be a critical force multiplier by facilitating the agency's enterprise management roles. In an enterprise transformed to provide assessment services rather than simply production, effective staff work would form the vital connective tissue between the assessment agency and its network of collectors and partners.

The force-multiplying quality of these staff functions will prove particularly valuable as agencies seek to manage both the demands of big data analytics and resource constraints. Assuming the U.S., Australian, and other ICs will continue to face tough budget and staffing pressures, any future investment in data analytics–related functions will likely come at the

expense of all-source analyst capacity, as analyst billets are retasked for new data-related missions. Investing more in staff functions would provide a scalable solution for the agency to leverage more external capacity to meet rising customer demands—and a scalable solution to maximize service delivery will become particularly salient in case of future budget or staffing cuts.

Thus, the future assessment agency should have a more diverse ecology of personnel. Rather than treating all-source analysts as the sole core competency and all other functions as ancillary support, an effective assessment agency that has adapted well to big data–related disruptive technologies will rely critically on the interaction of three core job types, none of which can be fully effective without the others: *data analytics disciplines*, including data scientists and software engineers, to process and manipulate big data inputs; *all-source analysts*, to provide expert and customized assessment advice; and *intelligence staff functions*, to manage and enable the assessment agency's key advantage: its connections to the customer and the rest of the enterprise.

Conclusion

Disruptive technologies carry implications not only for the work of the future analyst, but also for the future assessment agency. In particular, big data and its associated trends should yield enormous productivity and capability gains. But these technologies will also put pressure on the assessment agency as a whole to move away from internally producing all their intelligence and toward a service-provider model in which it tailors intelligence solutions sourced from across the IC and elsewhere. Many of these implications apply particularly to foundational military intelligence, so they will not be felt equally across the IC, and they will also extend to deployed warfighter support and collaboration with other government agencies and allied partners.

Like no change since the end of the Cold War, the advent of big data and data analytics will compel abiding changes in the IC. The risks and opportunities we have outlined are foreseeable in the next 5 to 10 years; other disruptive technologies not yet conceptualized (let alone fielded) will have other, unknowable effects in coming decades. The unknowable nature of future disruptive technologies, however, should not prevent IC leaders from executing a big data strategy immediately to transform both analysis and the enterprise.

None of these changes is inevitable; exploiting big data's remarkable opportunities and mitigating its risks demand strategic vision. An adaptive and effective defense intelligence enterprise will need new IT tools, new skill sets, and new business processes to embrace innovative technologies, and these will be costly. It will also entail a formidable recruitment and training challenge not only to cultivate a cadre of skilled data scientists but also to train all-source analysts on the uses and limits of data analytics. Meeting the challenge of big data will require investments of money and resources, and some risk-taking on new technologies and protocols—precisely at the moment of tightening budget constraints and post–Edward Snowden security sensitivities. These investments will have to compete with continued investments in the IC's treasured but exorbitant clandestine collection platforms, and IC leaders will need to make increasingly tough decisions on allocating those resources. As resources for traditional clandestine collection shrink, the obvious solution would be to reduce unnecessary duplication and dedicate those rare collection means to priority hard targets.

Most importantly, meeting the challenge of big data requires disciplined leadership to judge and maintain the right balance between data-intensive analytic functions, such as foundational defense intelligence, and making sense of complex phenomena for strategic intelligence advice. Absent strong direction, big data could easily become fetishized, where the quantity of data collected, collated, and processed becomes the measure of the community's effectiveness and distorts the analytic agenda. Instead, IC leadership must ensure that expertise and tradecraft are at the center of analytic operations and that knowledge creation and assessment services are at the center of enterprise management—all in the service, ultimately, of decision advantage for the customer. **JFQ**

Notes

[1] *Big data* is now a hackneyed, almost passé, term, but in the absence of a widely accepted substitute, it remains useful. For a non-scientific introduction to big data and its transformative potential, see Kenneth Neil Cukier and Viktor Mayer-Schoenberger, "The Rise of Big Data: How It's Changing the Way We Think about the World," *Foreign Affairs* (May–June 2013).

[2] James Manyika et al., *Disruptive Technologies: Advances that Will Transform Life, Business, and the Global Economy* (San Francisco: McKinsey Global Institute, May 2013), available at <www.mckinsey.com/insights/business_technology/disruptive_technologies>.

[3] On the skills required for data science, see Drew Conway, "Data Science in the U.S. Intelligence Community," *IQT Quarterly* 2, no. 4 (Spring 2011), 24–27. McKinsey estimates that by 2018 the demand for data-science talent will exceed its projected supply by about 50–60 percent (see Manyika et al.). The Intelligence Community will need to compete with the more lucrative private sector for those scarce talents.

[4] On object-based production and activity-based intelligence, see Catherine Johnston, "Modernizing Defense Intelligence: Object-Based Production and Activity-Based Intelligence," briefing, Defense Intelligence Agency, June 27, 2013, available at <www.ncsi.com/diaid/2013/presentations/johnston.pdf>.

[5] Kirk Borne, "Knowledge Discovery from Mining Big Data," briefing, March 12, 2013, available at <http://realserver4v.stsci.edu/t/data/2013/03/3194/KborneStsci2013.pdf>.

[6] On puzzles and mysteries, see Gregory F. Treverton, "Risks and Riddles," *Smithsonian Magazine* (June 2007).

[7] We are grateful to Josh Kerbel for coining this distinction between objects and phenomena.

[8] Josh Kerbel and Anthony Olcott, "The Intelligence-Policy Nexus: Synthesizing with Clients, Not Analyzing for Customers," *Studies in Intelligence* 54, no. 4 (December 2010).

[9] See, especially, Kerbell and Olcott; interview with Robert Blackwill, "A Policymaker's Perspective on Intelligence Analysis," *Studies in Intelligence* 38, no. 5 (1995); and Thomas Fingar, "Intelligence as a Service Industry," *The American Interest* 7, no. 4 (March–April 2012).