Lessons from SABLE SPEAR: The Application of an Artificial Intelligence Methodology in the Business of Intelligence

Craig A. Dudley

If in the other sciences we should arrive at certainty without doubt and truth without error, it behooves us to place the foundations of knowledge in mathematics.

—Roger Bacon

Project SABLE SPEAR, was a multiyear exploration into the opportunities and challenges of applying artificial intelligence (AI) fully into the intelligence process. The experiment provided insights into this new methodological approach to intelligence analysis. Standing in stark contrast to the intelligence methods that define current Intelligence Community (IC) analytic tradecraft, AI abstracts value in data and algorithms and centers original insights and the power of timely discovery in the open-source domain. This article explores the award-winning SABLE SPEAR journey and illuminates insights that will help to define how AI is applied within the IC and what will have to change in IC work if AI is employed.

❖ ❖ ❖

At the annual Defense Intelligence Agency (DIA) award ceremony in December 2019, Project SABLE SPEAR received a Team Award from the director of DIA. As I accepted the award on behalf of the team, the director said, "Of all the awards, this one intrigues me the most." I answered, "This is the future of our business," to which he replied, "I know."

The previous spring, Brian Drake, the leader of a team of all-source analysts working to understand the global flows of illicit fentanyl—one of the powerful synthetic opiods that cause tens of thousands of deaths each year—had come into my office at Joint Base Anacostia-Bolling with a proposal. DIA had funds available to invest in an "innovative idea" through the continuation of a relationship with a small Silicon Valley start-up that showed early success in applying AI to the production of finished intelligence. Brian's proposal was simple: although the start-up had built stability models based on historical data, he wanted to illuminate a complex, illicit network in its entirety as near to real time as possible. He would name the project SABLE SPEAR.

Brian's team had a typical cross section of intelligence analysts at various stages of careers in intelligence and with months of formal training in analytic tradecraft as prescribed in IC directives (ICD) such as ICD-203, "Analytic Standards," and ICD-206, "Sourcing Requirements for Disseminated Analytic Products."[a] Their formal training and the directives codified best practices in overcoming cognitive biases, avoiding

a. https://www.dni.gov/index.php/who-we-are/organizations/policy-capabilities/ps/ps-related-menus/ps-related-links/policy-division/intelligence-community-directives?high-

Although our analysts were experts in intelligence, we were certain we would struggle in the language of AI and requested NGA support interpreting between the two languages.

politicization, and communicating confidence in intelligence products.

We would begin to distinguish this method—elaborated in detail in academic works (including Mark Lowenthal's *Intelligence: From Secrets to Policy*[a]) and professional analytic tradecraft certification programs[b]—as "biological intelligence"a term used in the AI community to differentiate the typical analyst's process from the experience we were about to have with AI.

The team traveled to Palo Alto with two data scientists borrowed from the National Geospatial-Intelligence Agency (NGA). Although our analysts were experts in intelligence, we were certain we would struggle in the language of AI and requested NGA support interpreting between the two languages. Our initial discussions with the company included an overview of our intelligence problem—global trafficking in illicit fentanyl—and an overview of the company's approach to finding in big data environments associations between illicit behaviors and entities engaging in the behaviors.

The requirement we gave to the company was quite simple: *illuminate the networks associated with the distribution of illicit fentanyl.*

Before returning to Washington, we gave the vendor some of our understandings of the data sets that could be of particular value and some basic insights into patterns that characterized the phenomenon, but otherwise the company was limited entirely to the open-source domain and its original research. To enable effective auditing, the company was told to show its work to a level consistent with the analytic tradecraft standards used in citing evidence in finished intelligence. Drake's team would be available to provide guidance to the company and to validate the AI outputs.

Four months later the company sent representatives to Washington to present its initial findings. They were profound.

Across illicit entities and their associations, the company's outputs were numerically far superior to ours. The company's AI methods identified 100 percent more companies engaged in illicit activity, 400 percent more people so engaged, and counted 900 percent more illicit activities. In addition, the company's findings offered a "degree of fidelity we could not have anticipated."[c] Because the company had been told to "show its work," the empirical evidence used in drawing the characterizations and correlations were presented for examination and validation.

To be sure, some of the entities the vendor identified were deemed to be false positives by our analysts. That feedback was used to identify and correct the algorithmic framework that had falsely characterized the entities.

Most impressively though, the AI approach identified analytically relevant variables that our analysts probably would never have come up with and made instantaneous associations for those variables across multiple, often complex, data sets. Having identified the unique associative signatures for an "illicit actor" on the internet, for example, AI could then scan the entirety of the internet for that same associative pattern, illuminating considerably more entities within seconds.

Association, Intervention, and Imagination

The more we tried to understand and contextualize the AI outputs—and indeed find the words to explain the process clearly to our decision-makers—we found unique clarity in UCLA researcher Judea Pearl's work

light=WyJpbnRlbGxpZ2VuY2UgY29tbXVuaXR5IGRpcmVjdGl2ZXMiXQ==

a. Mark M. Lowenthal, *Intelligence: From Secrets to Policy* (CQ Press, 2000).

b. For example, the Department of Defense All Source Analysis Certification Program is part of the DoD-wide initiative to professionalize the intelligence field. The development of professional certification programs ensures an integrated, agile workforce that can meet the department's needs in a dynamic environment. Accessed 19 February 2020 at: https://dodcertp-mo.defense.gov/CDASA/

c. Brian Drake, DODIIS Worldwide Conference. Tampa, FL, 19 August 2019. Accessed at: https://www.dvidshub.net/video/703931/sable-spear-using-artificial-intelligence-confront-opioid-crisis

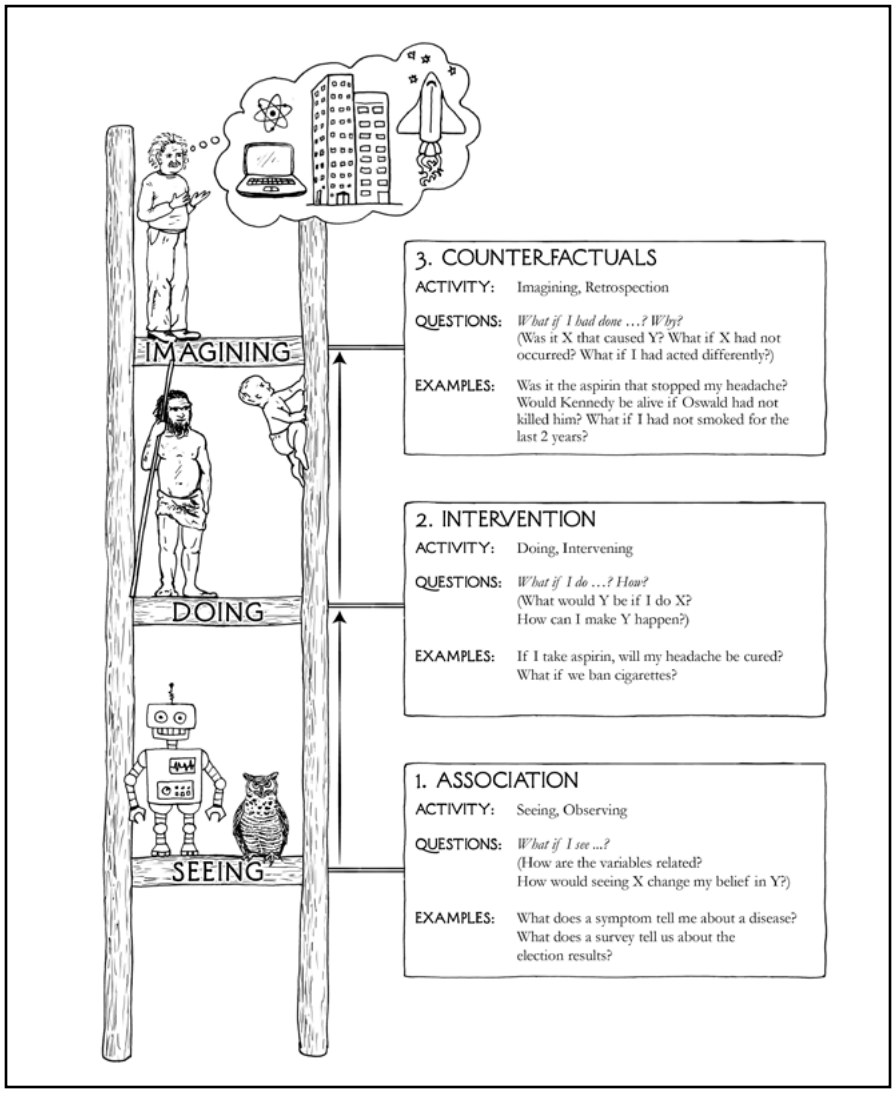
on the power of causal models.[a] The first rung in his Ladder of Causality calls for predictions based on "passive observation" and "characterized by the question 'what if I see…?'" What the AI team was providing us was the power of AI in this phase. In fact, according to Pearl, "Just as they did thirty years ago, machine learning programs (including those with deep neural networks) operate almost entirely in an associational mode."

In his *The Book of Why*, Pearl identifies advancements in causal science that were exactly what we began to experience in SABLE SPEAR. His "causal ladder" continues to help us to explain, in the business of intelligence, those analytic behaviors that can benefit immediately from AI (associations), the experiments that should now be pursued in the intervention phase, and the contributions that must continue to be served by human imagination.[b]

Aggregating and Presenting Data

As we began to refine the outputs from the associative phase, Brian's intelligence team validated the AI outputs and informed the development of a user interface that enabled the production of strategic intelligence and conveyed clarity and confidence in the empirical behaviors associated with individual entities. Aggregating and presenting the data allowed us to more accurately identify volumes of illicit fentanyl flows, major routes, and the entities commanding the greatest market share.

In fact, we soon had enough fidelity in associations to qualify an entity's relative criminal behavior in a "criminality index" as part of the trafficking ecosystem. The criminality index associated specific criminal behaviors as defined by criminal law—for example, association in a rapid and real-time process of the posting of an advertisement selling fentanyl with the entity (individual) making the post. In cases in which the volume of illicit behaviors an entity exhibited was higher (posting 30 ads versus posting four ads), the criminality score was elevated relative to other entities. Similarly, if an entity had a higher volume of "types" of US criminal code allegedly violated (selling fentanyl, and selling cocaine, and selling counterfeit documents) they would also have a higher criminality score.



"Ladder of Causality" © Maayan Visuals (http://www.maayanillustration.com/)

a. Judea Pearl, *The Book of Why: The New Science of Cause and Effect* (Basic Books, 2020).
b. Ibid., page 33.

Pearl argues that computers "cannot tell us what will happen in a counterfactual or imaginary world, in which some observed facts are bluntly negated.

By implication, this means that once the collective behavioral components of a given intelligence problem are resolved in an information environment, the collective associations that define issues like strategic missile deployment, the names and locations of intelligence officers, and the operational planning of extremist groups could be monitored and illuminated in near real time.

Protecting US Persons Information

We turned next to the issue of protecting information involving US persons. We told the vendor to assume every entity they encounter in the information environment is a US person and only after "proving they are not," through sufficient associations, could they be revealed to IC customers. For law enforcement customers these restrictions were not necessary.

We quickly found ourselves in an information environment where unique data holders—law enforcement entities at the federal, state, and local level, the Food and Drug Administration, the US Postal Inspection Service—each with an authority to intervene could do so more efficiently and more comprehensively by understanding the entire problem. Not only would these individual entities benefit from the sensemaking of their unique data, but they would benefit considerably from contextualizing their information holdings within the whole.

Issues of Intervention

Intervention is an area where we must continue to explore and invest in the development of causal models that allow for experimentation—to test the effects of "if we do this," what might happen as a result. According to Pearl, what is less widely known is that "successful predictions of the effects of interventions can sometimes be made even without an experiment. A sufficiently strong and accurate causal model can allow us to use rung-one (observational) data to answer rung-two (interventional) queries."[a]

Counterfactuals

Pearl argues that computers "cannot tell us what will happen in a counterfactual or imaginary world, in which some observed facts are bluntly negated. Yet the human mind does make such explanation-seeking inferences, reliably and repeatably." It is within this space that we recognize *the role of the all-source analyst will continue to be critical* — to contextualize the artificial outputs within the national security decision-making space we support as intelligence organizations. Consumers of intelligence will still need timely and comprehensive insights and the role of the all-source analysts in representing those outputs will continue to be central, even if the initial illumination of those insights is artificially derived.

Implications

Having used a grounded theory (GT) methodology in my doctoral research, I can attest that the methodological application created through this AI experiment was, in fact, analogous to GT, in which empirical phenomena are coded and then categorized for examination to develop "theoretical sampling" that explains themes within the data.[b] In strategic intelligence terms, this methodology achieved the same objectives as the investment in all-source analysts: the development of "foreknowledge"[c] (theoretical sampling).[d]

In the case of the AI method developed for SABLE SPEAR, this inductive GT approach happened rapidly and continuously, changing as quickly as the empirical underpinnings of the learned codes and categories; the derivative theoretical sampling (foreknowledge) was dynamic.

For strategic intelligence, foreknowledge could be achieved through AI that is inductive and constantly comparative, with dynamic developments in the information environment. As codes and categories

a. Ibid., 32.

b. Barney Glaser and Anselm Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research* (Aldine Publishing, 1967).

c. Used interchangeably here, foreknowledge and theoretical sampling both imply that future outcomes can to a degree be predictable; a theory is a coherent group of tested general propositions, commonly regarded as correct, that can be used as *principles of explanation and prediction for a class of phenomena.*

d. Theoretical sampling is a process of data collection for generating theory whereby the analyst jointly collects codes and data and decides what data to collect next and where to find them, in order to develop a theory as it is described in Barney Glaser, *Theoretical Sensitivity: Advances in the Methodology of Grounded Theory* (Sociology Press, 1978).

are identified and refined, the ecosystem moves closer to theoretical sampling (foreknowledge) at a pace far exceeding the human mind's cognitive limitations. While the purpose of IC directives— the timely and comprehensive representation of knowledge—would remain valid, the business model to get to that end-state would be more effective with AI.

In fact, the distinctive advantage of this approach may place Lowenthal's work and current intelligence doctrine cleanly in the annals of intelligence history.

For law enforcement, empirical phenomena in the information environment could be correlated instantly to federal, state, and local laws and the entities associated with the violation of those laws. In the second phase of SABLE SPEAR, we proved this scenario through our criminality index. As illicit entities enter and exit the information environment and their level of criminal behavior changes, so does their criminality score. Our use of the scoring system allowed for a prioritization of entities to be targeted, not for extensive investigation, but for validation and arrest.

Lessons in Applying AI

The SABLE SPEAR experiment taught us considerable lessons in the use of AI in our singular focus on a specific mission outcome: the illumination of illicit networks correlated to the marketing and distribution of one opioid. Through this process, a number of the experiences and challenges revealed details about the future of the intelligence business.

The application of AI, and the resources dedicated to that end, must begin with an expectation that the AI output is as timely and comprehensive as the outputs of the algorithms.

As our experience with AI deepened, we began to recognize the paradigmatic differences between the intelligence process of we humans and AI in the development of timely and comprehensive foreknowledge. In the case of our analysts, abstract value is in the minds of analysts, and the IC invests in training to improve expertise, logic, and argumentation, among other skills. Tradecraft, certifications, mentorship, and promotion frameworks are used to incentivize and reward these behaviors.

In the case of artificial intelligence, abstract value resides in data, algorithms, and the insights that can be derived from them. With data and algorithms taking center stage, conversations turn to defining the value of data sets and the level of effort and protocols needed to collect and protect those data.

Abstract Value Distinctions between Biological and Artificial Intelligence

The distinctions between the former and the latter intelligence must be understood as we evaluate technology for use within the Intelligence Community. Tools designed to assist all-source analysts to organize data, navigate cognitive obstacles, and illuminate correlations must be recognized as enabling the current biological intelligence process. In fact, the federal contractor market is saturated with vendors offering exactly these types of tools with varying levels of success.

The application of AI, and the resources dedicated to that end, must begin with an expectation that the AI output is only as timely and comprehensive as the outputs of the algorithms. These might include a real-time assessment of the likelihood of a strategic missile launch by an adversary, the real-time disposition of foreign intelligence officers, or the movement of illicit weapons among nefarious entities.

Ensuring the Provenance of Evidence

The need to "trace" the empirical correlations that form the foundations of an assessment can be algorithmically resolved within an AI ecosystem and tailored to the needs of contributing stakeholders. For example, if a law enforcement entity requires a standard of evidentiary integrity in judicial proceedings, pieces of evidence used to correlate an entity with criminal activity can be tailored into the production of "charge sheets" that manifest the data and their relationships to a degree sufficient to present in legal proceedings. Similarly, for the producers of strategic intelligence, the data can be adapted to meet to the analytical, argumentation, and presentation standards laid out in IC directives to serve policymaking at all levels.

Analysts' Roles Will Have to Change

All-source analysts, as generally known in the IC today, will differ from analysts who will be required to work with AI. Central to their new roles will be the application of yet-to-be-developed professional standards and processes by which analysts interact within the AI space.

The open-source environment is a common competitive space that must be the domain for the origination of comprehensive and timely discovery.

In addition, tradecraft certifications, IC directives prescribing standards and joint publications describing the roles of analysis in supporting warfighters must change. Integral to these guiding documents must be the articulation of where and how the power of AI will be leveraged to support intelligence customers.

In their new roles, analysts educate AI tools by prescribing the initial characterizations of the problem and assigning *initial relative value* to the data used for characterizing problems. Analysts must also serve the important role of validating the resulting outputs for their customers. As long as decisionmakers rely on cognitive processes, AI outputs must be presented in ways that allow decisionmakers to take advantage of their timeliness and comprehensiveness.

Similarly, the functions described in the common "intelligence cycle" take place simultaneously and in real time in the application of AI methods rather than as distinctive and sequential elements of collection management.

Leveraging Open Source

The open-source environment is a common competitive space that must be the domain for the origination of comprehensive and timely discovery. This is true for two reasons: first, the growing and disproportionate volume of analytically relevant data, for any issue, resides in the open-source domain. Second, the algorithmic environment, including new discoveries and relationships among algorithms, changes rapidly and continuously. *It is unreasonable to expect that the dynamic nature of the open source domain can be replicated in a classified environment and maintain the benefit of these phenomena.*

The ancient Greek philosopher Heraclitus is said to have observed that "no man ever steps in the same river twice, for it's not the same river and he's not the same man." Similarly, in the everchanging flow of data in the open-source domain, the data used for finding insight may be present one moment and gone the next. This reality is uniquely relevant when we consider moving unclassified data into a classified domain for analysis; there is a corresponding level of latency that affects decision advantage.

A helpful analogy we developed for characterizing the importance of open source was to compare it to the four center squares of a chess board. Holding and dominating the center enables more agile pieces of the enterprise (human intelligence, signals intelligence, etc.) to target information that cannot be discovered in publicly available information. In fact, the open-source domain takes center stage in defining what is and is not secret.

Redefining Data Ownership

For AI to work, data are *centrally valuable* to an assessment whether or not we are able to conceive of their relevance. To this end, *the mechanisms to protect an organization's unique data must reside in the algorithmic space and not be left to the judgment of individuals to determine what can and cannot be shared.*

One of the greatest obstacles to this end will be the sharing of data between intelligence and law enforcement organizations. While both communities have justifications for protecting the information they gather, their collective data must be accessible to a virtual AI environment in order to drastically improve the understanding of both entities and the collective. For example, if the US government is interested in addressing the opioid crisis, a comprehensive illumination of that problem means a detailed and real-time characterization of the problem in its entirety. To achieve that end state, AI must include all data from all agencies with responsibilities in that space, including the Drug Enforcement Administration, the Federal Bureau of Investigation, the US Postal Inspection Service, the Food and Drug Administration, and state and local governments where the most detailed consumer data exist.

Determining the Value of Sensitive Collection

Applying an AI method with origins in the open-source domain also means that agencies with a specific charter to collect information will have a mechanism to determine the relative value of that information based on its direct relationship to foreknowledge. For example, if an agency has the authority to collect signals or human intelligence, it will be able to *quantitatively* examine the value of that investment based on the weight of specific data points in advancing theoretical sampling. In today's intelligence framework, analysts are responsible for giving opinions on the value of data—a process that is plagued by shortfalls endemic to cognitive processes.

Knowing this, agencies will have to be ready to accept that specific collection programs may contribute surprisingly little to the resolution of intelligence problems or criminal investigations. Fortunately, the AI methodology will also facilitate an intelligent conversation about where unique collection capabilities need to be focused by defining what is truly unknown in the open information environment. It is in those areas that sensitive collection can be economically focused for a competitive advantage in decisionmaking.

What *is* well known within the IC is that considerable money is spent collecting information that can be known within the unclassified domain—things that are not really secrets.

Experimenting in the "Intervention" Space

A considerable advantage of applied AI is the ability to manipulate data algorithmically to test potential outcomes of actions before those actions take place. For example, in the characterization of an illicit network, an algorithmic modification can determine the effects of removing an entity from the network to determine the costs and benefits associated with that action. The derivative determination is repeatedly learned from previous instances within the information environment where a similar type of entity exited a similar type of network. This means the predicted effects are based on considerable volumes of data and activities rather than the few limited by human cognition.

More impressive, however, is that the machine could also recommend multiple and simultaneous, or sequential, actions to meet *defined objectives* within the AI environment. The AI ecosystem will be able to automatically generate a set of actions based on the objectives, constraints, and restraints of the analyst educating the ecosystem.

Agencies will have to be ready to accept that specific collection programs may contribute surprisingly little to the characterization of intelligence problems or criminal investigations.

Economic Efficiencies Inherent

Using AI to address national security issues would enable an exponential growth in the level of associations that can be developed across the whole of government, providing more courses of action for intervention. An agency's participation in an AI ecosystem would mean both the refined understanding of their organization's areas for action but also a considerable benefit to the collective as the data and users reach the critical mass needed to make it commercially attractive for data, tools, and expertise providers to feed their inputs into the ecosystem.

Commercial attractiveness requires that there be automated mechanisms in place that would make selling or providing data to the ecosystem rapid and painless for government and industry. Imagine how this would work in the absence of an ecosystem approach: the government would need to write contracts to purchase data only after a painfully slow requirements and procurement process. The process could take months, and what is worse, the information would most likely be irrelevant by the time it was made available.

Easing the process of data purchase by allowing ecosystem providers to make digital gateway mechanisms would transform today's slow data purchase process into a rapid commercial purchase between two commercial entities.

Once all of this data starts flowing into the ecosystem, it becomes automatically aggregated, connected, and curated in order to make the collective more useful for the entire community in an automatic and data policy managed way. The data policy manager would ensure that confidentiality, publicly identifiable information, and classification policies are strictly and conservatively adhered to.

Ultimately, the purpose behind incentivizing providers to input their data, tools, and expertise into the ecosystem is to have a multiplier effect on the number of associations that can be drawn between desired outputs and the variables available within the ecosystem. More associations will bring more possible points of interventions (what-if capabilities). More intervention points will provide more prescribed courses of action (guidance) for significantly changing the desired outcome.

Conclusion

The implications of applied AI are not evolutionary, but revolutionary, and would require investment changes to the tune of billions of dollars. It means the way intelligence and law enforcement conceptualize "intelligence" must radically change to include a new intelligence cycle in which an "analyst" serves to educate the initial development of an artificial ecosystem and the validation

The implications of applied AI are not evolutionary, but revolutionary, and would require investment changes to the tune of billions of dollars.

and communication of the artificially derived outputs. It means the types of people serving central roles in the intelligence business must be examined through their roles in the creation and interactions with artificial ecosystems.

The SABLE SPEAR experiment has allowed for an exploration of AI methods, but more such experiments are needed to fully understand the technical, human, policy, and legal requirements needed to effectively advance the business of intelligence. Each of these realities must continue to be debated, researched, and invested in to determine the types of people and resources needed to be competitive in the application of AI methods.

❖ ❖ ❖

The author: Craig A. Dudley is a division chief in the US Defense Intelligence Agency. During his 18-year career with the DIA, most of it overseas, he has had experience in capacity building, collection management, and all-source analysis, and served multiple tours at combatant commands and the Joint Staff. He holds a Doctorate in Comparative Intelligence from the University of Botswana, where his research focused on applied intelligence models.