# THE AIM INITIATIVE

## A STRATEGY FOR AUGMENTING INTELLIGENCE USING MACHINES

# Contents

This page intentionally left blank.

# THE AIM INITIATIVE

Augmenting Intelligence using Machines Increasing insight
and knowledge through Artificial Intelligence, Automation, and Augmentation

## Foreword

### FROM THE DIRECTOR OF NATIONAL INTELLIGENCE:

Closing the gap between decisions and data collection is a top priority for the Intelligence Community (IC). The pace at which data are generated and collected is increasing exponentially—and the IC workforce available to analyze and interpret this all-source, cross-domain data is not. Leveraging artificial intelligence, automation, and augmentation technologies to amplify the effectiveness of our workforce will advance mission capability and enhance the IC's ability to provide needed data interpretation to decision makers. The Augmenting Intelligence using Machines (AIM) Strategy provides the framework for the incorporation of AIM technologies to accelerate mission capability development across the IC. I challenge the IC workforce, based on the principles outlined in the AIM Strategy, to establish and implement an IC-wide AIM framework, inclusive of mission partners—be big but be practical—to provide real capability to close the gap between decisions being made and data collection.

Dan Coats
Director of National Intelligence

### FROM THE PRINCIPAL DEPUTY DIRECTOR OF NATIONAL INTELLIGENCE:

To meet its vision of ensuring intelligence advantage, the IC must adapt to the rapid global technological democratization in sensing, communications, computing, and machine analysis of data. These trends threaten to erode what were previously unique USIC capabilities and advantages; going forward, we must improve our ability to analyze and draw conclusions from IC-wide data collections at scale. I have identified AIM technologies as key transformative elements that will enable our analytic workforce to effectively leverage the increasing data volume for decision advantage. This document provides the overarching strategy and objectives for effective incorporation of AIM into the IC baseline. I welcome your feedback on this document.

Susan Gordon
Principal Deputy Director of National Intelligence

## Executive Summary

It is the job of the IC to analyze data, connect disparate data sets, apply context to data, infer meaning from data, and ultimately make analytic judgments based on all available data.  The pace at which data are generated, whether by collection or publically available information (PAI), is increasing exponentially and long ago exceeded our collective ability to understand it or to find the most relevant data with which to make analytic judgments.  AIM AAA technologies (Artificial intelligence, process Automation, and IC officer Augmentation) as key transformative elements are crucial for future mission success and efficiency.

This document outlines how the IC will incorporate AIM capabilities in a manner that resolves key IC legal, policy, cultural, technical, and structural challenges while producing optimally effective analytic and operational contributions to the intelligence mission.

Artificial intelligence (AI), especially its sub-discipline machine learning (ML), has shown dramatic advances in autonomous systems, computer vision, natural language processing, and game playing. These AI systems can perform tasks significantly beyond what was possible only recently (e.g., autonomous systems) and in some cases even beyond what humans can achieve (e.g., chess and Go).  In light of these recent advances, the IC is carefully considering methods for fully automating well-defined processes and augmenting human expertise with analytics or planning capabilities for their potential benefit. The IC is also monitoring these same technologies with respect to their vulnerabilities in development and adoption.  Accordingly, AIM seeks to determine how the IC can best manage uncertainty by achieving acceptable risk suited to the demonstrable analytic and operational advantages in AIM-enabled solutions and tradecraft.

Due to the widespread commercial application of these AI technologies, the private sector is making considerable investments in related infrastructure and people.  Therefore, we must carefully monitor and leverage private investment, focus our efforts on areas of unique mission need, and rethink how we attract and retain human expertise.  This strategic imperative exists because our adversaries, notably Russia and China, also recognize the potential for AI to transform military and intelligence operations and are investing aggressively to make that advantage a reality.

Individual components of the IC have already recognized the value of AAA technologies.  It is the goal of the AIM initiative to bring those disparate efforts together in order to maximize impact and accelerate development.  Increases in data volume and velocity are putting pressure on existing workflows, and our adversaries are putting significant effort into AI technologies that can blind or deceive the IC.  By adopting AIM, the IC will be able to meet those challenges.  This initiative leverages lessons learned from current and past AI efforts; strengthens the collaboration between the IC and industry, research agencies, and academic talent; and grows the talent pool of expertise for the IC.  We will continue to expand our interagency approach to AIM development to ensure that the implementation plan we deliver is the IC's plan as opposed to the Office of the Director of National Intelligence's (ODNI) AIM plan for the IC.

The AIM initiative will enable the IC to fundamentally change the way it produces intelligence.  We will achieve superiority by adopting the best available commercial AI applications and combining them with IC-unique algorithms and data holdings to augment the reasoning capabilities of our analysts. Simply stated, our goal is the following:

> **"If it is knowable, and it is important, then we know it."** – Sue Gordon

The AIM initiative is an IC-wide strategy for three reasons:

- First, there is intense competition in the private sector for AI and especially ML talent. The IC needs to establish new incentive and hiring models and stop competing internally for the same scarce resources.

- Second, AI and ML systems require large high-quality tagged data sets that must be shared with IC partners to the maximum extent allowable. Rule sets, algorithms, and expert knowledge bases that capture the tacit knowledge of intelligence domain experts must be available to all appropriate and relevant mission areas.

- Third, to rapidly accelerate AI adoption, the IC must have a solid digital foundation. This means leveraging the investment we have already made in the IC Information Technology Enterprise (IC ITE) and continuing to invest in and improve the IC ITE infrastructure.

The AIM initiative has four primary investment objectives:

**Objective 1 – Immediate and ongoing – Digital Foundation, Data, and Science and Technical Intelligence (S&TI):** AI activities are not a substitute for an enduring, secure, standardized, and measurable IC-wide digital infrastructure and data ecosystem; they are dependent on that foundation. In addition, the IC must improve foundational understanding of many aspects of AAA, to include a deeper understanding of the commercial supply chain, identification of ongoing developmental programs within the federal government that can be leveraged for a wider audience, and identification of adversarial uses of AI.

**Objective 2 – Short term – Adopt Commercial and Open Source Narrow AI Solutions:** The IC must leverage the existing private sector and government investments by rapidly transitioning the best available commercial and open source Narrow AI capabilities.

**Objective 3 – Medium term – Invest in the Gaps (AI Assurance and Multimodal AI):** To create and maintain strategic advantage, the IC must develop both the capability and capacity to take advantage of available data across all INTs and open source, and develop AI solutions that process and relate information from multiple modalities. To facilitate this, the IC must continue to implement policies to break down traditional INT stovepipes.

**Objective 4 – Long term – Invest in Basic Research Focused on Sense-Making:** It is not enough to simply fuse information from multiple modalities together in response to a single, narrow task. The construction of shared models is needed to provide the basis for trust between human and machine teams. This level of understanding demands basic research advances in representing knowledge; goals and intent; entity extraction from incomplete, multimodal data; and discourse generation.

Inclusive of all four objectives, it is critical for the IC to address issues of AI assurance, transparency, and reliability as well as potential adversarial uses of AI. The AIM initiative must include a continuous effort to both understand how AI algorithms may fail.

The AIM initiative is about much more than technology. Implementing the strategy will entail addressing workforce challenges and understanding and shaping the policies and authorities governing how the IC deploys and uses AI. The global nature of the challenge and the rapidly evolving technological and societal frameworks dictate that the IC must have strong partnerships with other government agencies, the private sector, foreign partners, national laboratories, Federally Funded Research and Development Centers (FFRDC), University Affiliated Research Centers (UARC), and academic institutions. Lastly, the AIM initiative includes a robust communication and outreach plan for the workforce, Congress, members of the Executive Branch, industry and foreign partners, and the American people.

This page intentionally left blank.

## Mission Imperative

The business of the IC, both in its raw material and its product, is intelligence, which comes from data. It is the job of the IC to analyze, connect, apply context, infer meaning, and ultimately make analytic and operational judgments based on all available data. The pace at which data are generated is increasing exponentially and is stressing our collective abilities. Some examples:

- By 2021, it is estimated that the data generated by global web traffic will reach 3.3ZB/year (up from 1.2ZB/year in 2016); this corresponds to 3.5 networked devices per global capita.[1]
- The Director of NGA has publically estimated that at the current, accelerating pace of collection, we would need over 8 million imagery analysts by 2037 to process all imagery data.[2]

One particular area of concern for the IC is related to AI mission assurance, especially in light of recent commercial efforts that utilize AI to generate high-quality, affordable forgeries of audio and video media. This could lead to widespread difficulties separating truth from fiction. Adding to this challenge is the problem that AI expertise is scarce, distributed around the world, and very limited in the IC.

## Overview

In addition to the vision and guiding principles, this document provides guidance on:

- Investments
- Partnering with industry, academia, research agencies, and national laboratories
- Creating a new policy framework and tradecraft expectation that enable AI and ML while simultaneously promoting safe use and mitigating risk
- Reshaping the IC workforce
- Engagement with the Department of Defense (DoD) and international partners
- A strategy for creating a classified activity to generate strategic advantage for years to come
- Research and development
- Governance and IC collaboration models
- A robust communications strategy for all of our constituents including Congress, the work force, our industry and international partners, and the American people
- Establishing consistent classification/declassification processes through AI that promote secure information sharing and facilitate appropriate transparency to the public

[1] "The Zettabyte Era: Trends and Analysis", June 2017. Cisco Report. https://www.cisco.com/c/en/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html

[2] Remarks from Director Robert Cardillo at 31st Annual Small Sats Conference, Utah, 8/7/2017 http://www.nga.mil/MediaRoom/speechesremarks/Pages/Small-Satellites/Small-Satellites---Big-Data.aspx

## Vision

The AIM initiative seeks to secure and maintain a strategic competitive information advantage for the IC through focused development and rapid adoption of AAA technologies.

The leading private sector companies, both in market capitalization and prospects for growth, all recognize the importance of digital infrastructure and have made massive and ongoing investments in related technologies such as cloud services and big data. Since each new generation of technology builds on the previous one, it is critical that the IC continue to invest in its digital foundations. This initiative will guide the IC to accelerate the adoption of digital and analytics transformation, identify mission use cases, build a coherent data ecosystem, acquire the appropriate AI tools, reshape the workforce, adapt new workflow processes, and change the culture.

The IC can and must do this. Our IC ITE investments in cloud technology and data services have paved the road to harness the power of unique data collections and insights to provide decision advantage at machine speed.

## Guiding Principles

The following guiding principles define the set of unwavering precepts that influence and guide the direction of the AIM strategy to facilitate cultural, political, and legal adoption across the IC.

**The opportunity is great; the threat is real; the approach must be bold:** Recognizing that strategic advantage is fleeting and fragile, the IC must be willing to rethink or abandon processes and mechanisms designed for an earlier era, establish disciplined engineering and operations practices, and maintain an absolute focus on assuring advantage in an intensely competitive global adversarial environment.

**ML models are IC assets:** Building on the IC ITE principle that "Data is an IC asset," machine learned models are also IC assets (as opposed to agency or INT-specific assets).

- Training and validation data sets: Most ML methods require large, high-quality, tagged data sets. These data sets are important IC assets and must be shared with IC partners to the maximum extent allowable.
- Rule sets, algorithms, and expert knowledge bases that capture the tacit knowledge of intelligence domain experts are also IC assets that must be shared with all appropriate and relevant mission areas.
- This community approach must also recognize and act on the need for INT-specific improvements, as they will be the main drivers for transformative capabilities in the near term.
- While improving INT-specific technologies, the community approach must take into account the need for correlated cross-INT data sets.
- Even when the actual training data cannot be shared, sharing the models derived from these training sets, along with the lessons learned from them, increases the value of these assets.

**AI can be a powerful tool, but we must recognize challenges:**

- ML classifiers are only as good as the data that is used to train them. For example, an image classifier that is trained with ground-based imagery may fail to classify images collected from overhead.

- Even state-of-the-art AI models are vulnerable to adversarial exploitation.

- AI and ML models are subject to "concept drift," i.e., the notion that in the real world data often arrives in streams and evolves over time in non-obvious ways. Therefore, the models must continually adapt to changes in the data environment so that opportunities to improve their accuracy are not missed.

- The IC should be aware of popular trends in AI but should stay focused on how we can best use the technology. When the media hype dies down, the IC must be ready to perform the long-term and difficult task of creating lasting operational value from these technologies.

- Many ML models do not include a description of their decision-making process in their standard output, and thus their results can be misunderstood by the casual user.

**AI assurance models and adversarial AI must be addressed in parallel with AI systems:** The level of effort to fool an AI algorithm is considerably lower than to develop them. Therefore:

- Intelligence systems must account for failure modes. For example, image classifiers may be fooled by very small changes in the input data, reinforcing the need for recurring human involvement in AI activities.[3]

- The IC must understand and anticipate how foreign entities may use AI and develop techniques and tactics to deny and disrupt those activities.

- The IC must develop intelligence systems that can demonstrate the underlying rationale behind decisions and responses to both users and overseers. For intelligence systems that make critical decisions regarding classification, dissemination, or life-critical decisions, such decisions and responses must be able to evince some degree of proof of correctness in addition to transparency.

- Recent developments in computer vision have resulted in approaches that can generate fake (altered or fabricated) images and audio recordings that are difficult to distinguish from unaltered digital media. The IC needs to develop ways of countering this capability.

**AI is not a substitute for developing a solid digital foundation; it requires that foundation:**

- The IC must continue to invest in and improve the IC ITE infrastructure and develop strategies for shared state-of-the-art hardware and/or other High Performance Computing (HPC) systems.

- The IC must accelerate activities that make data widely shareable. Need-to-know requirements and operational sensitivities will be honored but must not be used as an excuse to unnecessarily limit data sharing.

- The IC must create a sustained program of investment in the creation of high-quality training sets for the most important intelligence priorities.

[3] Su, J., Vargas, D., Sakurai, K., "One Pixel Attack for Fooling Deep Neural Networks", arXiv 1710.08864, Feb. 2018

Despite the perceived investment gap, the IC has opportunities:

- U.S. Government (USG) investment in AI is dwarfed by investment of the private sector, and the IC investment is a fraction of what Department of Defense (DoD) is investing.
- The IC must not only leverage the investment of the DoD and private sector, but we must also be prepared to invest in areas of unique interest to the IC.
- The IC should invest in areas critical to the IC mission where the private sector has few incentives to invest, such as low-shot learning and adversarial AI/AI assurance.

Common services are a priority, however, there is still a need for specialization:

- The IC must create common services for common capabilities in computer vision, human language technology, identity intelligence, process modeling, analytic discovery, automated planning, and other areas, while encouraging principled approaches to mission-specific specialization where appropriate.

Investments in the workforce must be made: The IC must develop a more technologically sophisticated and enterprise aware workforce. We must:

- Embrace strategic workforce planning and workforce analytics to address AAA workforce requirements and skill gaps.
- Invest in programs for training and retooling the existing workforce in skills essential to working in an AI-augmented environment.
- Redefine recruitment, compensation, and retention strategies to attract talent with high-demand skills.
- Develop and continually expand partnership programs with industry, including internship and externship programs, to increase the number of cleared individuals with relevant skills both in and out of government.
- Leverage the IC Joint Duty (JD) Program more strategically to share expertise across the IC in a seamless manner.
- Understand and maximize human capital authorities, policies, and programs to augment the AAA workforce.

Engagement with partners is essential:

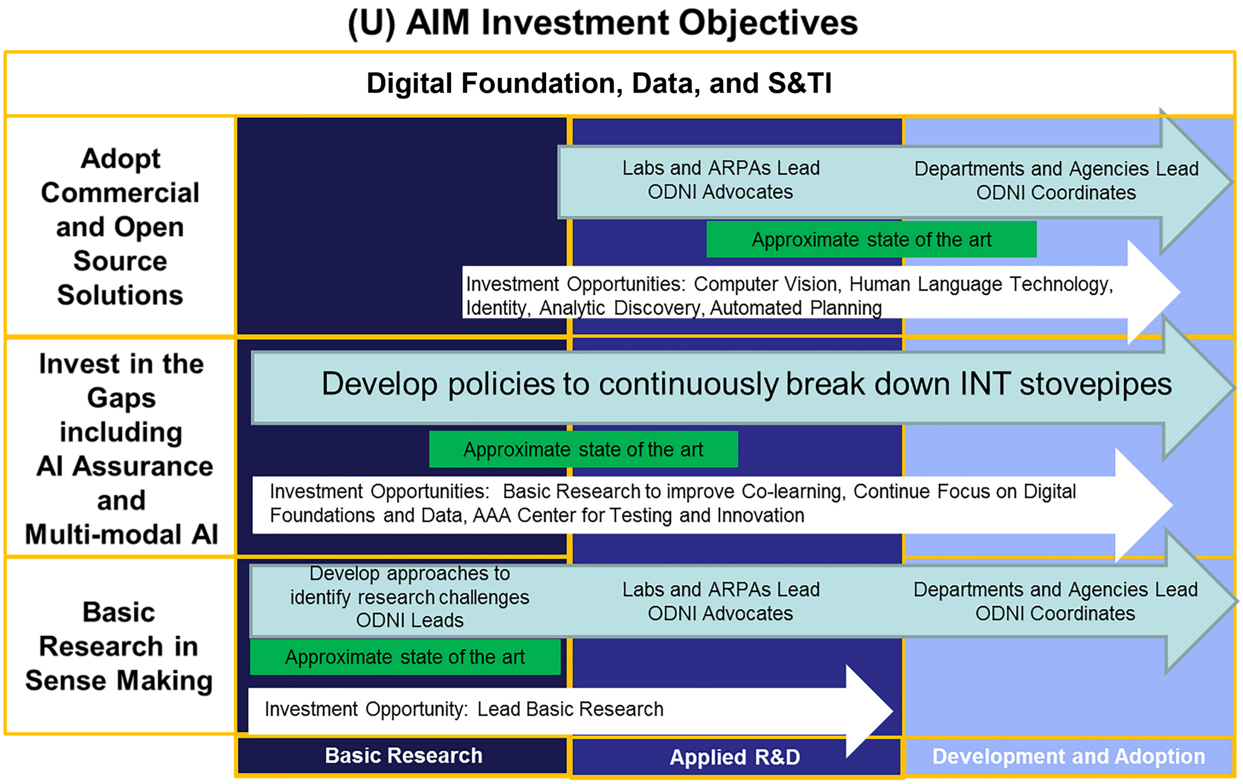
- A successful AI strategy requires engagement USG-wide, with the private sector, educational institutions, FFRDCs, national laboratories, and international partners (particularly Five Eye [FVEY] Partners).

Maintaining an understanding of the foreign threat is an intelligence priority:

- The IC must place an emphasis on S&TI integrated with operations and focused on AI in order to maintain strategic advantage, effectively counter these threats, and develop appropriate intelligence policies.

## Investment Strategy

Worldwide private sector investment in AI, ML, and related technologies is growing rapidly. Estimates of global private sector investment in 2016 range from $26B to $39B (McKinsey).[4] This investment strategy acknowledges the significant private sector investment and prioritizes investments that 1) allow the IC to rapidly adopt the best commercial and open source capabilities, and 2) accelerate research in those areas unique to the IC and where the private sector is not currently focused. A successful investment strategy also recognizes we must maintain momentum on foundational infrastructure gains, such as completing the IC's HPC architecture as well as accelerate data conditioning, storage, and sharing activities. This four-part investment plan, illustrated in Figure 1, addresses each aspect of basic research, applied R&D, and development and adoption.



(U) AIM Investment Objectives

Digital Foundation, Data, and S&TI

Adopt Commercial and Open Source Solutions

Labs and ARPAs Lead
ODNI Advocates

Departments and Agencies Lead
ODNI Coordinates

Approximate state of the art

Investment Opportunities: Computer Vision, Human Language Technology, Identity, Analytic Discovery, Automated Planning

Invest in the Gaps including AI Assurance and Multi-modal AI

Develop policies to continuously break down INT stovepipes

Approximate state of the art

Investment Opportunities: Basic Research to improve Co-learning, Continue Focus on Digital Foundations and Data, AAA Center for Testing and Innovation

Basic Research in Sense Making

Develop approaches to identify research challenges
ODNI Leads

Labs and ARPAs Lead
ODNI Advocates

Departments and Agencies Lead
ODNI Coordinates

Approximate state of the art

Investment Opportunity: Lead Basic Research

Basic Research

Applied R&D

Development and Adoption

Figure 1: AIM Investment Strategy

[4] "Artificial Intelligence – The Next Digital Frontier", June 2017, McKinsey Global Institute

**Objective 1 – Immediate and ongoing – Digital Foundation, Data, and S&TI:** AI activities are not a substitute for an enduring, secure, standardized, and measurable IC-wide digital infrastructure and data ecosystem. The IC will:

- Make data accessible to a wide variety of analytic platforms and models.
- Establish and maintain relevant training data across all INTs and disciplines.
- Adapt policies and tradecraft to enable more automated methods of assembling and vetting training data.
- Seek to future-proof data. Establish standards for data labeling and metrics for evaluation.
- Undertake a program of continuous growth in computational resources to ensure sufficient numbers of current generation hardware are available to IC AI practitioners.
- Improve foundational S&TI for AI, including adversarial uses of AI.
- Research co-learning and AI assurance models, especially vulnerabilities and standards.

**Objective 2 – Short term – Adopt Commercial and Open Source AAA Solutions:** The IC must leverage the massive private sector investment by rapidly transitioning the best available commercial and open source AAA capabilities. This will be accomplished as follows:

- **Aggressively pursue shovel-ready opportunities across IC Agencies**
- **Establish an IC AIM Center:** To foster innovation and rapidly prototype transformative solutions, the IC will establish an AIM Center staffed with AI and ML talent from across the IC, augmented by experts from industry.
- **Collaborate with key partners to identify opportunities:** Strengthen partnerships with the Intelligence Advanced Research Projects Activity (IARPA), the Defense Advanced Research Projects Agency, In-Q-Tel, the national laboratories, Defense Innovation Unit-Experimental, and industry. Advocate for those activities that address gaps with a minimum amount of duplicative effort, which will facilitate rapid transition of appropriate AAA capabilities to operations.

**Objective 3 – Medium term – AI Assurance and Multimodal AI:** In order to create and maintain strategic advantage, the IC must develop AI solutions that process and relate information from multiple modalities. To facilitate this, the IC must continue to implement policies that break down traditional INT stovepipes.

**Objective 4 – Long term – Invest in Basic Research Focused on Sense-Making:** The IC must understand multimodal information in context and look for ways that substantially augment the activities of IC Officers.

The massive investments to date as well as those currently planned (both public and private) will not be sufficient to meet the unique, foundational, and all-source sense-making needs of the IC. Therefore, basic research will focus on those areas and facilitate collaboration across the public and private sectors.

## Policy and Authorities

As part of the AIM strategy, the IC will examine the current tradecraft landscape and address emerging policy issues with appropriate efforts internal to the USG and, where needed, international venues. Policies to codify AIM activities (e.g., acquisition, enterprise management, classification, and analytic integrity) will be developed with consultation from appropriate general counsel, civil liberties, privacy, and policy personnel. ODNI will provide a dedicated, integrated policy and legal effort to break down barriers to information sharing, particularly INT-specific data, so that we do not inadvertently slow the pace of technological progress.

## Workforce Strategy

The IC must develop a more technologically sophisticated and enterprise-aware workforce. We must:

**Embrace strategic workforce planning and analytics:** Workforce planning will aid in accurately identifying current and future skill gaps, and will also enhance the IC's ability to determine the most appropriate mitigation strategies (e.g., training, compensation, etc.).

**Invest in programs for training and equipping the workforce in essential AI skills:** This does not mean everyone in the workforce needs to become an expert in deep learning or Python coding, but everyone does need to understand how AIM fits into the new workflow and how they can contribute. Specific actions include:

- Leadership – must understand the implications on the intelligence process, have a sophisticated understanding of the threat environment and foster an environment that enables an open and collaborative culture while reskilling the workforce to operate in an AI accelerated environment.

**Build on the IC 2025 workforce transformation to attract talent with high-demand AI skills:** The role descriptions for people with these skills have gone by many different terms in recent years. Therefore, individuals with these skills may be available but under different keywords. These alternative terms include analytics, data science, data wrangling, statistics, ML, deep learning, and modeling. These cover both the researchers who propose and test new methods, as well as model builders who use these algorithms to create and validate models.

**Develop partnership programs with industry and academia** to increase the pool of people inside the IC with awareness of best practices and available tools in this fast moving area, and to encourage individuals outside of the IC to build capabilities that meet the needs of the IC. Specific actions include:

- Recruit talent before graduation, and before competition with industry salaries, through service-for-education agreements ("ROTC"), expansion of IC postdocs, and internship/externship programs.
- Support temporary non-government to government (internship, externship) rotations.
- Expand sabbaticals, part-time industry Intergovernmental Personnel Act positions, and consultancies that grant clearances to faculty to increase the available technical skill available to the IC.

- Investigate changes in policy or funding to improve retention and attraction of U.S. national and foreign-born graduates in technical fields, including staff roles that do not require a clearance, and "fast-track" hiring that allows experts to perform productive work before obtaining a clearance.
- Expand use of open challenge problems (e.g., IARPA) and develop data and proxy problems that focus external communities on IC regions of interest.
- Identify unclassified equivalent domains for researchers to pursue. While the IC represents a unique environment, often similar domains give uncleared researchers an opportunity to develop and test algorithms on data that has many of the same qualities as IC data. This also fosters an interest in public service.

**Leverage the IC JD Program:** As competition for talent continues to increase outside of the IC, the community must leverage the IC Civilian JD program to share and retain talent across the IC and provide the workforce opportunities in other IC missions. We should:

- Identify related positions in each agency that will benefit from the JD program.
- Track JD opportunities for professionals and the AI community's use of the JD program.
- Ensure that the return on investment of personnel participating in the JD program meets AI objectives and is sustainable through the sense-making investment stage.

**Understand and maximize human capital authorities and policies to augment the AI workforce:** IC elements and the DNI have certain authorities at their disposal to assist in the management of the IC employment lifecycle. In order to ensure the most effective use of these authorities, we should:

- Identify and implement authorities that will create efficiencies in recruiting, hiring, compensation, training, and retention of AI professionals.
- Ensure that human capital policies enable IC elements to support the employment of AI personnel and do not erect barriers that may disengage the AI workforce.

**Leverage current human capital programs and monitor implementation and user feedback:** AI managers must continually collaborate with human capital professionals to take advantage of programs that enable the IC workforce to meet mission objectives. Examples include:

- Scholarships and other educational financial aid (e.g., Stokes Scholarships).
- Well-rounded recruiting programs that include outreach to diverse schools (e.g., Adopt-A-School, IC Wounded Warrior Program, STEM Outreach).
- IC Heritage Community Liaison Council, which is a forum that supports IC workforce development objectives, including outreach and recruitment.
- Recruiting efforts such as the IC Virtual Career Fair and IC Centers for Academic Excellence.

## Industry Partnership Strategy

Since the bulk of the nation's AI resources reside in the private sector, partnership is essential to the IC. Yet the barriers working with government often require considerable effort to clear. This requires a more flexible acquisition paradigm. This includes cooperative agreements that may trade data for algorithms or "Analysis-as-a-Service," as well as public prize challenges to solve IC problems. With the bulk of development occurring outside the IC, we must collectively prioritize Certification and Accreditation of new software so that code can more quickly be deployed on secure networks. ODNI, in collaboration with the IC elements, will develop an industry partnership plan for AIM capabilities. Elements of the plan will include:

- Industry access to USG data for algorithm development
- Enabling human resource strategies to simplify the development and sharing of AI skills between government and industry to include new approaches to security
- ODNI advocacy for AI basic research funding
- Creating AI services of common concern or specific capability contracts
- Update intelligence and industry data- and capability-sharing policies and oversight

## Roles for USG Agencies, National Labs, FFRDC, UARC, Commercial and Academic Institutions

To capitalize on the combined capabilities of the USG, national laboratories, private industry, and academic institutions, the ODNI must facilitate partner integration. Therefore, partner roles include:

IC:

- Promote communications between AIM partners
- Promote development of shared analytic services where feasible
- Share datasets and computing
- Capture and share expert knowledge from IC systems and analytics
- Capture and share mission data for future training datasets and simulations
- Develop defensive and offensive techniques for adversarial/counter AI
- Coordinate R&D activities
- Modernize multi-agency data sharing practices
- Improve S&TI on foreign AIM capabilities and intentions

Whole-of-USG:

- Share datasets across labs, private industry, and academic institutions
- Acquire and retain experts on immigration policy, IPAs, or service-for-education agreements ("ROTC")
- Coordinate DoD and IC R&D, computing and data purchases, and data-labeling efforts
- Synchronize funding for basic and applied research efforts

National Labs/FFRDC/UARC

- Provide expert advisors to USG
- Verify and validate algorithms and data sets, testing and evaluation (T&E), and AI methodology
- Conduct mission-focused research
- Develop AIM-related algorithm and systems prototyping
- Support talent pipeline development

Industry

- Provide commercial tools accessible through USG acquisitions and/or investment
- Conduct mission-focused, AIM-related research and development
- Provide expert advisors to USG
- Appropriately share datasets through a legal, supportable business model

Academic Institutions

- Perform the research needed to develop long-term scientific breakthroughs
- Provide expert advisors to USG
- Train the next generation to be a highly skilled workforce equipped to develop AAA tools and develop skills to utilize AAA systems

## Five Eye Foreign Partner Engagement

Allied and partner nations can enhance our joint development of intelligence products. Expanding international partnerships will provide opportunities to increase collection access and reliability, improve the quality and quantity of partner data and analysis, align strategic capabilities and emerging technologies, and promote compatibility across digital architectures and analytic tradecraft.

## AI Assurance – Secure and Maintain Competitive Advantage

The unique data and tools that the IC creates using those data are important IC assets that provide competitive advantage for USG missions. That advantage is fleeting and fragile, requiring disciplined engineering and operations practices, and an absolute focus on assuring advantage in an intensely competitive global adversarial environment. Commercial and USG needs differ in important ways but largely overlap with the concomitant requirement for continuous investment in data, tradecraft, tools, T&E, security, and S&TI.

AI technologies have clearly demonstrated that they can provide powerful capabilities. They have also demonstrated their brittleness and vulnerabilities. There are some principles and best practices that can be used today.

- **Data:** ML systems are only as good as the data used to train them. Acquiring those data in volume from the intended operational environments is a critical advantage. These data must be continuously monitored and reacquired as necessary. This is an engineering tradecraft best practice, akin to standard software test suite discipline.

- **Software:** The leading edge AI/ML software suites were written to support science, not national security operations. There is no notion of cyber security. USG needs are not aligned exactly with those of industry and universities; we need to differentiate in the state-of-the-evolving-art tools in a robust, sustainable way.

- **Systems:** Continuous evaluation of performance is required. There is very little theory to inform us as to when ML systems fail, or even whether they will work as expected[5]. This situation is not acceptable for any safety-critical or national security system. We must always incorporate performance monitoring, and we should support theory development.

- **Test and Evaluation:** Too many AI/ML projects launch without metrics to allow the IC to understand whether the investment is on track to succeed or fail. Create the discipline to define metrics up front and establish rigorous testing regimes and schedules.

**Concept Drift must be addressed.** "Concept Drift" is the idea that all computer tools are built with specific assumptions about the real world and that the basis for these assumptions generally changes over time, requiring the tools be monitored and updated. Best practices in established disciplines such as control systems theory can help structure how this challenge is attacked; we must detect issues and— when possible—automatically correct.

**Adversarial AI techniques represent opportunities and risks.** We have highly sophisticated adversaries with access to the same tools, their own data, and experts trained in the same universities as our own people. AI is merely one of the new battlegrounds for a technology-based arms race.

**S&TI is a priority.** We must develop a better understanding of foreign adversary tactics, techniques, and procedures.

[5] Papernot, N., McDaniel, P., Jha, S., Fredriskson, M., Celik, Z., Swami, A., "The Limitations of Deep Learning in Adversarial Settings", IEEE European Symposium on Security and Privacy, IEEE 2016, Saarbrucken, Germany

**Understanding when AAA techniques fail is critical.** The technical literature is replete with examples of how to deceive AAA systems.[6] We need to know how and where adversarial systems are in use against our assets.

## Outreach / Communications Strategy

A key factor in the success of transformation efforts like the AIM initiative comes through awareness and education of all of the varied constituents of the enterprise. Therefore, the ODNI will establish and maintain a robust communications engagement strategy for each of the following audiences:

- The IC, including leadership and the workforce
- The DoD and other government agencies
- Congress and the White House
- The private sector
- The national laboratories and academia
- The American people

## Governance

Following the example of private sector firms that are successfully implementing AI and recognizing that strong executive leadership goes hand in hand with stronger AI adoption, the PDDNI will, along with the IC Deputy Executive Committee (DEXCOM), serve as the executive sponsors for the strategy.

## Conclusion

AIM technologies will have a transformative effect on how the IC operates. Increases in data volumes and velocity require the IC to dramatically rethink how we perform our mission. Additionally, our adversaries have recognized the importance of AIM methods and are putting significant effort into these technologies. The principles and strategies laid out here will allow us to meet those challenges. Most notably, those strategies will build on and leverage lessons learned from current and successful AIM efforts; strengthen the collaboration between the IC and industry, research agencies, and academic talent; and grow the talent pool of AIM technology expertise for the IC. Our goal in all of this is to meet our IC objective now and into the future. "If it is knowable, and it is important, then we know it." – Sue Gordon

[6] Goodfellow, Ian, "Attacking Machine Learning with Adversarial Examples", https://blog.openai.com/adversarial-example-research/

## Appendix A: Background on AI

Artificial Intelligence (AI): The IC defines AI as "the branch of computer science focused on programming machines to perform tasks that replicate or augment aspects of human cognition," a term coined in the 1950s.[7] At that time, scientists began to harness nascent computer capabilities to perform advanced information manipulations much more rapidly. In particular, it was realized that computers could be used not only to perform calculations on numbers, but also to perform inference on other types of information such as symbols, data, and text. This popularized the idea of a "thinking machine" that could, if filled with all the right knowledge and rules for access and retrieval, simulate a human response.[8]

Technologies and research areas generally considered to be sub-domains to AI:

- Automated Planning and Scheduling
- Computer Vision
- Decision Support, Predictive Analytics, and Analytic Discovery
- Distributed Artificial Intelligence/Agent-based Systems
- Human Language Technologies
- Identity Intelligence
- ML
- Process Modeling
- Robotics/Autonomous Systems

Ideal AI System: A machine capable of ideal human intelligence with a computer's speed, capacity, and precision.[9]

Adversarial AI: A subset of AI focused on understanding how AI systems behave in the presence of a malicious adversary.

Artificial Narrow Intelligence (ANI): Also known as "Narrow AI" or "weak" AI, this is an AI system that is specialized for a single purpose and cannot be generalized. All current applications are ANIs.

Artificial General Intelligence (AGI): Also known as "General AI" or "strong" AI, this is an AI system that can handle any human intellectual task—memory, learning, abstraction, and creativity. There are no AGI systems in existence, although building an AGI has been the goal of the field since it was founded in the 1950s.

The AIM INITIATIVE—Augmenting Intelligence using Machines

Narrow AI and Multimodal AI: Nearly all current commercial applications of AI are narrow solutions in that they solve a single problem with a single kind of data. Image classification, face recognition, and human language translation are all examples of narrow AI solutions. The IC must bring together

[7] National Intelligence Council Sense of the Community Memorandum SOCM 2016-039C, 24 June 2016

[8] Fiscal Years 2019-2030 Major Issue Study Final Report, Advanced Analytics, Deep Learning, and Artificial Intelligence.

[9] Yost, Kirk, "Threats Posed by Advances in Artificial Intelligence", MITRE Technical Report sponsored by OSD Office of Net Assessment, 2016

data from multiple INTs to provide context and meaning to analysts over a variety of different data. Multimodal AI presents a whole new group of challenges in a number of areas that the IC must overcome. The challenges include:

- Representation - Presenting and summarizing multimodal data in a way that exploits its complementarity and redundancy. For example, development of representations that allow simultaneous analysis of audio derived from SIGINT with imagery and video.
- Translation – Learning how to translate or map one mode to another while recognizing that the relationship between modalities is often subjective. For example, there are any number of correct ways to describe an image with words, but a perfect translation from image to text may not exist.
- Alignment – Understanding how to identify direct relationships between elements and sub-elements to derive meaning from multiple modalities. For example, aligning a verbal description of an event with sequences in a video requires measuring similarity between modalities and understanding long-range dependencies and ambiguity.
- Fusion – Understanding how to join information from multiple modalities, which may have different predictive power and noise characteristics. For example, in audio-visual speech recognition, the visual description of the lip motion is fused with the speech signal to predict spoken words.
- Co-learning – Exploring how knowledge gained learning from one modality can help computational models trained on a different modality.

Automated Planning: A branch of AI focused on generating strategies or action sequences necessary to achieve a goal.

Automation: Computational systems designed to perform repetitive tasks.

Autonomous Systems: Systems that carry out tasks without human intervention. In AIM we are especially focused on computational systems that perform complex reasoning tasks.

Catastrophic Forgetting: A learning problem which occurs when performance learned in earlier tasks in a series is entirely or mostly lost after being given examples of later tasks.

Co-learning: A sub area of machine learning focused on either understanding how multiple agents can simultaneously learn, or how a single agent can use learning from one modality to improve computational models trained on a different modality.

Computer Vision: A field of study that aims to analyze, extract, and understand objects and relationships from within single or multiple images.

Concept Drift: The notion in ML that the concept being learned will change over time. Differentiating between different types of malware, for example, is a classification task that changes as new malware is produced.

Deep Learning: “Representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. With the composition of enough such transformations, very complex functions can be learned.” (LeCun, Y., Bengio, Y., and Hinton, G., “Deep Learning”, Nature, Vol 521, 2015.)

Graphical Processing Unit (GPU): Specialized electronics designed to perform rapid mathematical functions to render images, animations, and videos.

Human or Intelligence Augmentation: Use of information technology to augment human intelligence in the performance of some task. Unlike autonomous systems which aim to replace human activity, augmentation is designed with humans as central.

Knowledge Discovery: A process of discovering useful knowledge from a collection of data.

Low-shot Learning: An object recognition, ML classification task where learning must take place despite having only one, or a few, example images for training.

Machine Learning: The field of study interested in building computational systems that can improve their own performance of some task.

Machine Learning Classifier: A ML model designed to assign given examples into known discrete categories (i.e., classification).

Machine Learning Model: An explicit summary of data which is useful for performing some task. The product of ML systems like decision tree algorithms or neural networks are generically known as models.

Multimodal AI: A subset of AI focused on methods that emphasize the integration of linguistic, acoustic, and visual data in the completion of some task.

Natural Language Processing: A field of study that aims to analyze and understand human language communications both spoken and textual. Can include analysis and generation of language.

Sense-making: A process of creating understanding in situations of high complexity.

Technical Debt: Complications accumulated during the construction and use of software or ML models that make maintenance of these models difficult (e.g., hidden feedback sources, undeclared consumers, data dependencies, and changes in the external world).

Testing data: A collection of examples used to evaluate the performance of a ML Model.

Training data: A collection of examples used to inform the construction of a ML model.