DEPARTMENT OF DEFENSE ARTIFICIAL INTELLIGENCE, BIG DATA AND CLOUD TAXONOMY

Foreword by Hon. Robert O. Work, 32nd Deputy Secretary of Defense



COMPANIES INCLUDED

Altamira Technologies Corp. Aptima Inc. AT&T Inc. (T) BAE Systems PLC (BAESY) Boeing Co. (BA) Booz Allen Hamilton Inc. (BAH) CACI International Inc. (CACI) Carahsoft Technology Corp. Combined Technical Services LLC Concurrent Technologies Corp. Cray Inc. CSRA Inc. (CSRA) deciBel Research Inc. Dell Inc. Deloitte Consulting LLP DLT Solutions Inc. **DSD** Laboratories DXC Technologies Corp. (DXC) General Dynamics Corp. (GD) Georgia Tech. Harris Corp. Johns Hopkins APL HRL Laboratories International Business Machines Inc. (IBM) Inforeliance Corp. Insight Public Sector Inc. Intelligent Automation Inc. Intelligent Software Solutions Inc. JF Taylor Inc. L3 Technologies Inc. (LLL) Leidos Inc. (LDOS) Lockheed Martin Co. (LMT) Logos Technologies LLC ManTech International Corp. (MANT) Microsoft Corp. (MSFT) Mythics Inc. NANA Regional Corp. Northrop Grumman Corp. (NOC) Orbital ATK Inc. Product Data Integration Technologies Inc. QinetiQ PLC **Ravenswood Solutions** Raytheon Co. (RTN) RDM Corp. Regents of University of California Rockwell Collins Inc. (COL) SAIC Corp. (SAIC) SAP SE (SAP) Sierra Nevada Corp. Silicon Graphics International Corp. SK Telecom Co. Soar Technology Inc. SRC Inc. SRI International Inc. Telos Corp. (TLSRP) Teradata Corp. (TDC) Tetra Tech Inc. (TTEK) Torch Technologies University of Pennsylvania

University of Southern California United Technologies Corp. Vector Planning & Services Inc. Vision Systems Inc. Vykin Corp. World Wide Technology Inc.

GOVINI CLIENT MARKET VIEWS

Artificial Intelligence - Computer Vision Artificial Intelligence - Data Mining Artificial Intelligence - Deep Learning Artificial Intelligence - Machine Learning Artificial Intelligence - Modeling & Simulation Artificial Intelligence - Natural Language Processing Artificial Intelligence - Neuromorphic Engineering Artificial Intelligence - Quantum Computing Artificial Intelligence - Supercomputing Artificial Intelligence - Virtual Agents Artificial Intelligence - Virtual Reality Big Data - Business Analytics Big Data - Data Analytics Big Data - Data Architecture & Modeling Big Data - Data Collection Big Data - Data Hygiene Big Data - Data Hygiene Software Big Data - Data Visualization Software Big Data - Data Warehouse Big Data - Distributed Processing Software Big Data - ETL & Data Processing Big Data - Intelligence Exploitation Cloud - IaaS Cloud - PaaS Cloud - SaaS

DEPARTMENT OF DEFENSE ARTIFICIAL INTELLIGENCE, BIG DATA AND CLOUD STANDARD MARKET TAXONOMY



Exhibit 1: Govini's Artificial Intelligence (AI) Big Data, and Cloud Taxonomy consists of three broad categories (gray boxes) associated with mission; seven segments (orange boxes) that define capabilities and 25 subsegments (white boxes) that constitute technological approach. The hierarchical organizational structure is designed to deliver insight ranging from high-level spending trends to granular details on specific programs and technical solutions. Current year spending and the five-year compound annual growth rate (CAGR) from FY2012 through FY2017 are noted for each sub-segment. Final FY2017 values are estimated based on public spending data that were available through October 2017.

Mining

\$48.0M

.govini

ARTIFICIAL INTELLIGENCE, AUTONOMOUS SYSTEMS AND THE THIRD OFFSET

Hon. Robert O. Work, 32nd Deputy Secretary of Defense

Autonomy results from delegation of decision to an authorized entity to take action within specific boundaries. An important distinction is that systems governed by prescriptive rules that permit no deviations are automated, but are not autonomous. To be autonomous, a system must have the capability to independently compose and select among different courses of action to accomplish assigned goals based on its knowledge and understanding of the world, itself and the situation.¹

Rapid advances in Artificial Intelligence (AI)—and the vastly improved autonomous systems and operations they will enable—are pointing towards new and more novel warfighting applications involving human-machine collaboration and combat teaming. These new applications will be the primary drivers of an emerging military-technical revolution. Military revolutions "are periods of discontinuous change that render obsolete or subordinate existing means for conducting war."² The U.S. military can either lead the coming revolution, or fall victim to it.

This stark choice will be determined by the degree to which the Department of Defense (DoD) recognizes the revolutionary military potential of AI and advanced autonomous systems; ramps up research and development in their associated technologies, such as advanced computing, artificial neural networks, computer vision, natural language processing, big data, machine learning, and unmanned systems and robotics; and aggressively develops the new systems, operational concepts and organizational constructs that exploit them in warfare.

The "Third Offset Strategy," first articulated by Secretary of Defense Chuck Hagel in November, 2014, announced DoD's intentions to lead the coming AI/autonomy driven military-technical revolution. By exploiting advances in AI and autonomous systems to improve the warfighting potential and performance of the U.S. military, the Strategy aims to restore the Joint Force's eroding conventional overmatch versus any potential adversary, thereby strengthening conventional deterrence.

As its name suggests, the Third Offset Strategy follows two previous competitive strategies with similar ends. DoD adopted the Second Offset Strategy in the mid to late-1970s to overcome the Warsaw Pact's large numerical advantage in conventional forces along the Central European front. Up until then, the U.S. had offset superior Warsaw Pact numbers with a smaller force armed with battlefield atomic weapons. However, this First Offset Strategy was blunted once the Soviets achieved strategic nuclear parity with the U.S., which called into question NATO's threat to employ tactical nuclear weapons in its defense. A new strategy was needed to strengthen conventional deterrence. Rather than try to match the communist military machine soldier-for-soldier, tank-for-tank, or plane-for-plane, U.S. planners once again opted to offset Soviet strength—this time by developing an ability to "look deep and shoot deep" and destroy follow-on Warsaw Pact forces before they reached NATO front lines. Looking deep and shooting deep required the development of a new and far more capable theater-wide battle



network able to target and attack advancing Warsaw Pact forces still far away from the "forward edge of the battle area." Battle networks were nothing new. They first appeared at the start of World War II in the form of the British homeland air defense system, with the four interconnected grids that defined all subsequent battle networks:

- A sensor grid capable of wide area surveillance as well as narrower battlefield reconnaissance and targeting;
- A C4I grid (Command, Control, Communications, Computer and Intelligence Grid), able to make sense of what was happening in the area of operations, facilitate decisions to seek combat advantage, and transmit orders to the...
- ...effects grid, consisting of a wide array of kinetic and non-kinetic combat forces and effectors designed to achieve specific battlefield outcomes; and
- A sustainment and regeneration grid designed to sustain combat operations and regenerate combat losses.

What made the new NATO Follow on Forces Attack (FOFA) network so powerful was the marriage of long-range sensors and a new generation of guided munitions and submunitions, linked by digitally enabled, real-time battle management capabilities. Take for example the TR-1 aircraft, a modification of the iconic U-2 spy plane. Carrying side-looking radar at high altitudes, it could image Warsaw Pact armored forces operating over 100 miles from the NATO front lines. It would then downlink its data directly to ground processing centers which then quickly sent accurate firing data to the Army Tactical Missile System (ATACMS), a ballistic missile armed with guided submunitions. Similarly, the Joint Surveillance and Target Attack Radar System (JSTARS), with its ground moving target indicator mode and onboard battle management capabilities, could vector NATO tactical fighter-bombers armed with a variety of guided munitions towards deep Warsaw Pact armor formations. At the theater level, the Air Force created what is now known as the Combined Air Operations Center, or CAOC, which would issue coordinated air tasking orders to all NATO air forces.

Happily, NATO and Warsaw Pact forces never came to blows. However, the guided munitionsbattle network revolution spurred by the Second Offset Strategy was on display for all to see during Operation Desert Storm—the 1990-91 U.S.-led campaign to eject Iraqi forces from Kuwait. Although it was quite short, this campaign clearly demonstrated that the combination of theater-wide sensor grids, digital C4I grids designed for real-time battle management, and effects grids that emphasize guided weapon attacks rendered subordinate combined arms warfare characterized by massed formations employing unguided weapons fire.³

Because the U.S. military was the aggressive first mover in guided munitions-battle network warfare, it enjoyed a dominant conventional overmatch versus any regional competitor in the immediate post-Cold War period. Now, however, newly emerging great power competitors like Russia and China are rapidly achieving parity in guided munitions-battle network warfare, and Second Offset technologies are proliferating worldwide. As a result, the conventional overmatch enjoyed by the Joint Force over the past two-and-a-half decades is now eroding—and at an accelerating rate. This circumstance is challenging traditional means of U.S. power projection and undermining conventional deterrence, thereby raising the future risk of conventional interstate warfare.



Hence the need for a Third Offset Strategy, which seeks to exploit advances in AI and autonomous systems to improve the performance of Joint Force guided munitions-battle networks in five different ways:

- Deep learning machines, powered by artificial neural networks and trained with big data sets, and inserted in every battle network grid;
- New ways of human-machine collaboration, which rely on AI-enabled learning machines to help humans make more timely and relevant combat decisions;
- New ways to facilitate assisted human operations, whereby smart AI devices will allow operators of all types can plug into and call upon the power of the entire Joint Force battle network to accomplish assigned missions and tasks;
- New types of human-machine combat teaming that see seamless coordinated operations between manned and unmanned systems, including those that are increasingly autonomous in their operations; and
- Cyber and electronic warfare-hardened network-enabled, autonomous and high-speed weapons capable of collaborative attacks.

By gradually reconfiguring Joint Force battle networks in these ways, and by adopting new operational concepts and organizational constructs to exploit them, Joint Force battle networks will be able to sense and perceive battlefield patterns more readily and rapidly, facilitate more timely and relevant combat decisions, and apply more rapid, discreet and accurate effects with less loss of life. If all these things happen, the Joint Force will operate at a higher, more effective tempo than its adversaries, and thereby gain an important, if not decisive, advantage in both campaign and tactical level operations.

Importantly, however, the Third Offset Strategy recognizes that much of the research and development of AI and autonomous systems is being conducted in the commercial sector, meaning its fruits are available to all competitors. This means the competition to derive advantage from them will be very intense. For example, China, has similar ambitions with respect to these important technologies. It is betting on AI to drive its future military and economic strengths—so much so that its AI strategy calls for China to become the world leader in the field by 2030. Only a similarly focused effort will keep the U.S. Joint Force from falling behind, and being on the wrong end of a new generation of human-machine warfare.

The Third Offset Strategy's desirable ends are well within reach if tied to an urgent and concerted DoD initiative to pursue them. This initiative must consist of two complementary efforts. The first is a robust, focused, and prioritized research and development (R&D) program designed to explore and field the technologies necessary to implement the strategy's vision. This is the realm of the Office of the Secretary of Defense (OSD), specifically the new Undersecretary of Defense for Research and Engineering. The second is the development of new operational concepts and organizational constructs to exploit these new technologies. This effort will fall primarily to the military services and combatant commands to implement.

.govini

The purpose of the following report, prepared by Matt Hummer and the staff at Govini, is to support the OSD effort. Govini's powerful data gathering and analytic tools can track Federal contract spending down to the penny. Using these tools, this report displays and analyzes actual DoD spending on Third Offset technologies from FY2012 through FY2017. By so doing, it shows whether or not the Department of Defense is "putting its money where its mouth is" on the Third Offset Strategy.

When reading the report, please keep in mind three things:

- First, this report is not an argument about the pros and cons of autonomous systems. It is a summation of the DoD spending on the technologies that support their development, fielding and operations. It is aimed primarily at DoD and national level decision makers who must decide if the amount and focus of that spending is appropriate.
- Second, the following data reflects spending only on unclassified contracts. One can expect there
 to be additional Third Offset spending on classified contracts, so this report tells only part of the
 story. Nevertheless, spending on unclassified contracts provides a clear indicator for the overall
 seriousness of DoD Third Offset efforts.
- Third, the Govini Strategic Intelligence Platform allows for the natural aggregation of spending data into "market taxonomies." As will soon be apparent, the initial market taxonomy separates current DoD Third Offset spending into three broad technology categories: artificial intelligence; big data technologies needed for machine learning; and the cloud services needed to store big data sets. While this taxonomy is illuminating, the new Under Secretary of Defense for Research and Engineering may find a different taxonomy more useful, and Govini's Strategic Intelligence Platform allows for the flexible aggregation and display of data in ways most helpful to the user. For the time being, however, this first report provides a good baseline from which the new Under Secretary can build a prioritized AI/Autonomy R&D portfolio in support of the Third Offset Strategy.

¹ L.G. Shattuck, Transitioning to Autonomy: A Human Systems Integration Perspective, as cited in the Report of the Defense Science Board Summer Study on Autonomy, June 2016, p.4.

² Michael G. Vickers and Robert C. Martinage, The Revolution in War (Washington, D.C.: Center for Strategic and Budgetary Assessments, December 2004), pp. 2-3.

³ Those who study military-technical revolutions or revolutions in military affairs consider Desert Storm the "defining battle" of the guided munitions-battle network revolution. A defining battle occurs when one of the forces involved demonstrates the dominance of a new way of war. Ibid., p. 3.





AI, BIG DATA AND CLOUD: CORNERSTONES OF THE THIRD OFFSET STRATEGY IN HIGHLY CONTESTED DOMAINS

DoD is investing in a wide array of technologies to enhance its military edge over large state competitors such as China and Russia. Of these, none are more important than Artificial Intelligence (AI) and Autonomous Systems, which are the technological cornerstones of the Department's Third Offset Strategy. As explained by Bob Work, by exploiting advances in AI and Autonomous Systems to improve the warfighting potential and performance of the U.S. military, the Strategy aims to restore the Joint Force's eroding conventional overmatch versus any potential adversary, thereby strengthening conventional deterrence.

This analytic report leverages data science to present Govini's AI, Big Data and Cloud Taxonomy, a roadmap for tracking major drivers of the these emerging technologies. The Taxonomy consists of three broad categories of spending: Artificial Intelligence, Big Data, and Cloud. Each of these categories are further divided into segments that define associated capabilities and subsegments that constitute a technological approach. The hierarchical organizational structure is designed to deliver insight ranging from high-level spending trends to granular details on specific programs and technical solutions over the last six fiscal years as means for predicting budget priorities in FY2019 and beyond.

The Taxonomy shows sub-segments within the the Artificial Intelligence category having the most growth since FY2012. All together, AI spending grew by a CAGR of 14.5 percent. Cloud spending also increased by a CAGR of 8.9 percent and Big Data spending increased by 0.7 percent.

Cloud



FY17 Contract Obligations Compared to 5 YR CAGR by Sub-Segment

Big Data

AI Big Data Taxonomy Categories

Artificial Intelligence

Exhibit 2: AI sub-segments (orange) had the most grow from FY2012 through FY2017. Spending on Computer Vision grew the most by 19 percent to \$492.3 million in FY2017. Virtual Agents, another AI sub-segment grew by a CAGR of 16.9 percent, followed by Quantum Computing, Natural Language Processing, Deep Learning, Machine Learning, Modeling & Simulation and Neuromorphic Engineering.



Key Findings

- DoD spending on AI, Big Data and Cloud reached \$7.4 billion in FY2017, which is 32.4 percent higher than the \$5.6 billion spent in FY2012. In FY2017, AI accounted for 33 percent of the spending total, while Big Data accounted for 47.9 percent and Cloud accounted for 19.1 percent.
- While AI accounted for only 33 percent of the FY2017 total, it contributed significantly to the overall growth in spending from FY2012. DoD spending in the three AI segments—Learning & Intelligence, Advanced Computing and AI Systems—grew the most from FY2012 through FY2017 by CAGRs of 13.7 percent, 11.6 percent and 16.4 percent, respectively.
- Within the Learning & Intelligence segment, Natural Language Processing spending grew the most by 16.8 percent to \$82.9 million in FY2017. Deep Learning spending increased the second most by 14.9 percent to \$238.1 million in FY2017 and Machine Learning followed with an increase of 14.3 percent to \$195.5 million in FY2017.
- Within the Advanced Computing segment, Quantum Computing spending increased the most by a CAGR of 16.8 percent to reach \$68.8 million in FY2017. Neuromorphic Engineering followed with an increase of 13.7 percent, reaching \$97.8 million in FY2017. Supercomputing, the largest and most mature sub-segment, had CAGR of 9.7 percent with spending reaching \$258.4 million in FY2017.
- Within the AI Systems segment, Computer Vision spending grew the most by a CAGR of 19 percent from FY2012 through FY2017. Virtual Agents, the smallest AI Systems sub-segment, grew by 16.9 percent to \$98.3 million in FY2017. Spending in the largest sub-segment, Virtual Reality grew by 14.3 percent to \$557.5 million in FY2017.
- Most Computer Vision investment is related to Intelligence, Surveillance and Reconnaissance (ISR), specifically advancing image definition, widening field-of-view and micro chips for processing imagery data. Virtual Reality spending is primarily for battle simulation and training. The least mature AI System, Virtual Agents, is focused on large-scale language processing and translation and is being funded primarily through Defense Advanced Research Programs Agency (DARPA) programs.
- While advancing AI has been a priority of DoD, other investments in foundational categories such as Big Data and Cloud must be made in order for AI to reach its full potential. This is especially true of Big Data, which is central to "teaching" learning machines. When rolling up aggregate spending, Big Data—consisting of technologies and services for collecting, processing and analyzing data—represents the largest category, accounting for 54.4 percent of total spending from FY2012 through FY2017.
- Big Data segments had modest spending growth compared to segments in other categories. Spending on Analytics, the largest Taxonomy segment by contract obligations, grew by a CAGR of 0.5 percent. Big Data Technologies spending grew by 0.7 percent and Data Collection & Processing spending grew by a CAGR of 1.4 percent.
- The most efficient way to facilitate access to Big Data is to store it in the Cloud. Cloud is currently the smallest of the three spending categories; it grew by a CAGR of 8.9 percent reaching a high of \$1.4 million in FY2017. However, with the recent announcement that DoD is accelerating a shift to the Cloud, this number is likely to rise.



DEPARTMENT OF DEFENSE ARTIFICIAL INTELLIGENCE STANDARD MARKET TAXONOMY



Govini's *AI, Big Data and Cloud Taxonomy* shows that AI is gaining traction within DoD by being integrated with operating concepts of mission systems. AI Systems was not only one of the largest Taxonomy segments by obligations, but also had the most spending growth. It accounted for 14.9 percent of overall Taxonomy spending since FY2012 and had a CAGR of 16.4 percent. The other AI-related segments also had strong spending growth. Learning & Intelligence spending grew by a CAGR of 13.7 percent and Advanced Computing grew by a CAGR of 11.6 percent.

DoD is Investing Heavily to Advance AI Learning & Intelligence Capabilities and Tools

Learning & Intelligence spending by DoD increased by a CAGR of 13.7 percent, the second most of all the *AI, Big Data and Cloud Taxonomy* segments. The overall growth in spending should not come as a surprise, but the prioritization of spending across sub-segments should.

Natural Language Processing (NLP) spending grew the most with a CAGR of 16.8 percent. DARPA fueled the spending growth accounting for 60.5 percent of the six-year total, through marquee programs such as Broad Operational Language Translation (BOLT) and Low Resources Languages for Emergency Incidents (LORELEI).

Other Learning & Intelligence sub-segments also had strong spending growth. Deep Learning spending grew by a CAGR of 14.9 percent and Machine Learning spending grew by 14.3 percent.

Govini has categorized Learning & Intelligence into the following five sub-segments:

- Modeling & Simulation facilitating understanding of system behavior without testing
- Deep Learning mimicking cognitive functions such as learning or problem solving
- Machine Learning the ability for computers to learn without being explicitly programmed
- Natural Language Processing programming to process large natural language corpa
- Data Mining discovering patterns in large data sets and transforming the data into understandable structures for further analysis

Learning & Intelligence Sub-Segments Modeling & Simulation Machine Learning

Modeling & Simulation
 Deep Learning

Data Mining

earning Natural Language Processing

FY17 Contract Obligations Compared to 5 YR CAGR by Sub-Segment



Exhibit 3: Annual spending increased significantly in most sub-segments with the exception of Data Mining. Natural Language Processing spending increased the most by 16.8 percent, followed by Deep Learning with a 14.9 percent spending increase and Machine Learning with a 14.3 percent spending increase. DARPA funded 60.5 percent of NLP programs and 28.9 percent of Deep Learning programs.



System Integrators are Seeking to Acquire Advanced Learning & Intelligence Capabilities

FY2016 marked a rapid rise in Learning & Intelligence spending. Modeling & Simulation, Deep Learning, and Data Mining spending increased the most by dollar value. Whereas FY2017 proved to have slightly different priorities such as Machine Learning and Natural Language Processing.

The fact remains that all fields of Learning & Intelligence theories are important to advancing DoD AI capabilities. They are also broadly deployed across programs funded by several DoD agencies including Army, Air Force Life Cycle Management Center (AFLCMC) and Missile Defense Agency (MDA).

Spending in the largest sub-segment, Modeling & Simulation is spread across the services with Navy and Army taking the lead through their warfare analyses and sensor simulation programs. Leidos, SAIC, AECOM and Orbital ATK led capture of Navy program spending and SAIC, CACI, Torch Technologies and Millennium Engineering led capture of Army spending. Northrop Grumman captured the most spending by MDA through its ballistic missile defense system threat software modeling work. Raytheon followed capturing 19.5 percent of MDA spending.

Each of these integrators are seeking advanced capabilities in Learning & Intelligence that will help differentiate their competitive offering. Northrop Grumman, Raytheon, Lockheed Martin and SAIC are all well represented across the other four sub-segments, but less so in the Deep Learning sub-segment. Most of the market leaders are not well established integrators rather little-known companies such as Aptima, Intelligence Automation, Soar Technology and Decibel Research. Machine Learning is similar, with Intelligent Software Solutions as the market leader through its WebTAS platform that integrates and visualizes data from multi-source data.

Learning & Intelligence Sub-Segments Modeling & Simulation Machine Learning Data Mining





Annual Contract Obligations by Sub-Segment

Vendor Market Share



Free sub-segments, Modeling & Simulation, Deep Learning and Machine Learning accounted for 82.6 percent of total segment spending since FY2012. However, spending in smaller sub-segments like Natural Language Processing grew the most by 16.8 percent since FY2012.



Advanced Computing Allows AI to Expand Beyond Narrow System Applications

Advanced computing power is the real enabler of AI. It allows machines to figure out how to perform tasks after being exposed to learning algorithms and training data.

DoD continues to invest in advancing computing capabilities, especially in the most recent fiscal year. Overall segment spending increased by 85.9 percent to \$424.9 million in FY2017 from \$228.5 million in FY2016. DARPA accounted for 37.3 percent of spending since FY2012, the most of all DoD funding offices. Naval Sea Systems Command, Army Program Office for Simulation, Training and Instrumentation (PEO STRI) and Warner Robins Air Logistics Center are also among the largest spenders on Advanced Computing.

Govini has categorized Advanced Computing into the following three sub-segments:

- Supercomputing compute performance measured in floating-point operations per second (FLOPS)
- Neuromorphic Engineering use of very-large-scale integration (VLSI) systems containing electronic analog circuits to mimic neuro-biological architectures
- Quantum Computing use of quantum bits (qbits), which can be in superpositions of states instead of binary bits, which is always in one or two definite states (0 or 1)

Advanced Computing Sub-Segments

Supercomputing

Neuromorphic Engineering
 Qua

Quantum Computing





Exhibit 5: The largest Advanced Computing sub-segment, Supercomputing grew by a CAGR of 9.7 percent to \$258.4 million in FY2017. The smaller sub-segments, Quantum Computing and Neuromorphic Engineering had stronger growth of 16.8 percent and 13.7 percent respectively. Spending on Neuromorphic Engineering reached \$97.8 million in FY2017 and Quantum Computing spending reached \$68.8 million. DARPA accounted for 37.3 percent of overall segment spending followed by Air Force Life Cycle Management Center and Office of Naval Research (ONR).



Advanced Computing Spending Increased Sharply in FY2017 and Will Continue to Grow

Ten years ago, spending on advanced computing was rationalized mostly by scientific leadership. Today, there are more practical reasons for spending; advanced computers underpin Deep Learning and Autonomy. All of which have cross-cutting applications within DoD operational concepts.

Such is the purpose of the Third Offset Strategy and AI and Advanced Computing are linchpins to successful implementation of the Strategy. Supercomputing, the largest and most mature sub-segment, is dominated by Cray, which accounted for 25.1 percent of direct capture. Cray's computers are also used by several other contractors as part of their technical solutions. IBM is also a big player in the defense market along with several others including Nvidia, Asetek, Aspen Systems, Gidel and Atipa.

The other sub-segments, Neuromorphic Engineering and Quantum Computing are less mature than Supercomputing as evident by type of organizations performing and funding contracts. DARPA accounts for 53.8 percent of spending on Neuromorphic Engineering of which large portions were obligated to Regents of University of California, IBM, University of Southern California (USC), Massachusetts Institute of Technology (MIT) and HRL Laboratories. Some of the same organizations are performing Quantum Computing contracts, notably HRL Laboratories and USC. Much like Neuromorphic Engineering, DARPA is funding most of the Quantum Computing work, accounting for 68.2 percent of sub-segment spending since FY2012.

Quantum Computing

Advanced Computing Sub-Segments

Supercomputing



Annual Contract Obligations by Sub-Segment

Neuromorphic Engineering

Vendor Market Share



Exhibit 6: Supercomputing, the largest sub-segment, accounted for 51.3 percent of overall Advanced Computing spending since FY2012. The sub-segment's dominance can be explained by a spike in FY2017 spending. Neuromorphic Engineering spending also increased significantly in FY2017 by 26.3 percent. Army led spending on Supercomputing, while DARPA led spending in the other sub-segments.



AI Systems Segment Spending Grows Most of All AI Taxonomy Segments Since FY2012

DoD has already begun to integrate AI with mission systems and operating concepts. While the applications are narrowly defined, several years of spending increases provide indication that AI has gained traction moving beyond test and development phase.

Virtual Reality (VR), the largest sub-segment by contract obligations, is the most mature given the sustained high levels of investment. As the sustained funding suggests, VR is redefining planning, simulation, and training across battle domains.

Computer Vision, another mature AI Systems sub-segment is also gaining traction. Sub-segment spending increased by 19 percent, the most of any Taxonomy sub-segment. Each service is spending to advance the capability. Army is investing in high resolution 3D geospatial information and Air Force is spending on several capabilities including advanced synthetic airborne radar sensors, while Navy sees promise for Computer Vision in multi-spectral targeting.

Virtual Agents, although the smallest sub-segment by contract obligations, is attracting investment. Spending grew by 16.9 percent between FY2012 and FY2017 with a large portion being allocated by DARPA

Govini has defined AI Systems in the following three sub-segments:

- Virtual Reality environments which provide a virtual presence and artificial affects
- Computer Vision systems that automate human vision tasks, including acquiring, processing and analyzing digital images and high-dimensional data
- Virtual Agents abstract functional systems that respond to a wide array of questions

AI Systems Sub-Segments Computer Vision

Virtual Reality

Virtual Agents

FY17 Contract Obligations Compared to 5 YR CAGR by Sub-Segment



Exhibit 7: Computer Vision, the second largest sub-segment by contract obligations had the most spending growth of 19 percent from FY2012 through FY2017. Virtual Agents grew by 16.9 percent and Virtual Reality grew by 14.3 percent.



AI Has Moved Beyond R&D and is Beginning to Play Strong Role in Mission Systems

AI has great potential for creating asymmetric advantages in warfare. Its speed and accuracy in reacting, adapting and predicting scenarios makes it the cornerstone of DoD's Third Offset Strategy. While there are several challenges to widespread adoption of AI, DoD has begun to invest in applications where AI can match human cognition for specific purposes.

Virtual Reality for battle simulation and training is an example of one of these applications. AI has proven in several situations to match and even outperform the very best human cognition. As a result, DoD has embraced this particular application of AI as a solution. So much so that leading providers of virtualized simulation and training including Raytheon, Lockheed Martin, SRI International and JF Taylor are seeking to integrate AI with their solutions if they have not already.

Computer Vision, the second largest AI System sub-segment, is being applied in many missions critical to warfighting. The most obvious is Intelligence Surveillance and Reconnaissance (ISR) where Leidos supports Army's Geospatial Center's High Resolution 3-D Geospatial Information Program and Raytheon delivers multi-spectral targeting systems among other Computer Vision-related technologies. Research and development areas include high-resolution, wide-field-of-view gigapixel cameras and neuromorphic microchips for processing imagery data. Still other applications of Computer Vision are being funded including detection of defects to aircraft and undersea terrain mapping.

AI Systems Sub-Segments

```
Virtual Reality
```



Virtual Agents

Annual Contract Obligations by Sub-Segment

Computer Vision

Vendor Market Share



Exhibit 8: Virtual Reality and Computer Vision are the most mature and largest sub-segments. The two accounted for 93.3 percent of AI Systems segment spending since FY2012. Investment in the smallest sub-segment, Virtual Agents comes primarily from DARPA and Navy Research Labs.



DEPARTMENT OF DEFENSE BIG DATA STANDARD MARKET TAXONOMY



Govini's *AI, Big Data and Cloud Taxonomy* present the Big Data category as the largest by obligated contracts accounting for 54.4 percent of overall spending since FY2012. Big Data is foundational to Artificial Intelligence. Data Collection and Processing is used for amassing the large quantities of high-fidelity data required to train machines. Big Data Technologies such as Data Warehouse, Distributed Processing Software and Data Architecture & Modeling are used to manipulate data for real-time processing and pattern recognition. Analytics is used to surface insight critical to human-machine teaming.

DoD has Prioritized Data Processing and Data Hygiene Over Data Collection

AI can only be as smart as the data ingested, which is one reason why DoD spending on Data Processing and Hygiene grew the most of all Data Collection & Processing sub-segments.

Data quality is an often overlooked challenge in most Big Data projects. It has the potential to compromise results and lead to misinformed decision making. Effectively querying heterogeneous data sources, then extracting, transforming and loading data towards one or more data models also greatly impacts overall data quality.

Nonetheless, collecting data is not enough. AI simply will not work without data that is properly standardized, normalized, de-duplicated, verified and enriched, with verifying and enriching among the most critical steps for making data useful. Without AI, the Third Offset Strategy will fall well short of its intended objective of widening the military capability gap between the U.S. and potential adversaries and strengthening conventional deterrence.

Govini has categorized Data Collection & Processing into the following four sub-segments:

- Data Collection the process of gathering information in a systematic fashion
- Extraction Transformation & Loading (ETL) & Data Processing three functions used to pull data out of staging databases and place them into production databases
- Data Hygiene the process to ensure that data is free from error and in a usable format
- Data Hygiene Software software that detects and corrects corrupt or inaccurate records

Data Collection & Processing Sub-Segments

ETL & Data Processing
 Data Collection

Data Hygiene
 Data Hygiene Software

FY17 Contract Obligations Compared to 5 YR CAGR by Sub-Segment



t 9: ETL & Data Processing became the largest sub-segment by FY2017 obligations accounting for 35.6 percent of total segment spending. Its emergence was fueled by a high annual spending growth rate of 8.3 percent. Other sub-segments related to data quality and usability, Data Hygiene and Data Hygiene Software, also had significant spending growth of 5.2 percent and 3.9 percent respectively.



Integrators Hold the Keys to Unlocking Potential Analytic Value of AI and Big Data

The digital universe is growing faster than ever. The world is expected to produce 44 zettabytes by 2020 and 163 zettabytes of data by 2025; only 4.4 zettabytes were produced in 2013. Video and images makes up a large portion of the digital data and this is especially true for DoD as it works to implement the Third Offset Strategy.

Despite the deluge of data, only a fraction of it has been explored for analytic value. By 2020, it is estimated that only 33 percent of the digital universe will contain information that has analytic value. In a world overwhelmed by information, Data Processing and Data Hygiene are critical to determining the value of data and unlocking the potential of AI and Big Data.

In the most recent fiscal year, DoD prioritized ETL & Data Processing, with spending increasing by 87.2 percent to \$243.6 million. A large portion of the increased FY2017 spending was captured by systems integrators including General Dynamics, Raytheon, EHR Total Solutions, Booz Allen Hamilton, GeoNorth Information Systems and Johns Hopkins Applied Physics Lab.

Data Hygiene spending also increased significantly in FY2017 by 53.6 percent to \$151.5 million. Lockheed Martin and Leidos led the market accounting for a combined 32.5 percent of total revenue captured. Northrop Grumman and Raytheon followed accounting for 6.5 percent and 4.8 percent of FY2017 spending respectively.

Data Collection & Processing Sub-Segments

Data Collection
 ETL & Data Processing

Data Hygiene
Data Hygiene Software

Annual Contract Obligations by Sub-Segment



Vendor Market Share



Exhibit 10: Systems integrators and technical service providers lead a Data Collection and Processing market more focused on data processing and cleansing than collection. Raytheon leads overall segment market share with 5.2 percent, mostly from its presence in ETL & Data Processing and Data Collection. Leidos, General Dynamics, Northrop Grumman, Lockheed Martin, DXC Technology, QinetiQ and CACI also rank among the top ten providers of Data Collection & Processing. Contractors that have a presence across Data Collection & Processing sub-segments have a competitive advantage.



The Convergence of Big Data and AI Set to Create Immense Value for DoD

Although many AI technologies have been in existence for several decades, only now are they able to take advantage of datasets of sufficient size to provide meaningful learning and results. Much of the credit goes to Big Data Technologies; without them easy access to large volumes of data and the ability scale ingestion would not be possible.

Advancement in Big Data technologies has certainly been helpful to AI and more broadly the Third Offset Strategy. But in the future, it may be AI that helps Big Data technologies to progress further, leading to the automation of decision making along logic trees made possible by Big Data and AI working together.

Govini has categorized Big Data Technologies into the following three sub-segments:

- Data Warehouse repository of integrated data from one or more disparate sources, which are routinely manipulated and processed
- Distributed Processing Software software used to manage shared resources in data processing, standardization and normalization
- Data Architecture & Modeling collection of policies, models, rules and standards that govern which data is collected and how it is stored, arranged, integrated and put into data architectures and systems

Big Data Technologies Sub-Segments

Data Warehouse Distributed Processing Software Data Architecture & Modeling

FY17 Contract Obligations Compared to 5 YR CAGR by Sub-Segment



Exhibit 11: DoD is prioritizing investment in Data Architecture & Modeling and Distributed Processing Software over Data Warehouse. Annual spending on Data Architecture & Modeling grew by a CAGR of 6.8 percent to \$69 million in FY2017. Spending on Distributed Processing Software grew by a CAGR of 2.7 percent to \$71.3 million. Annual spending on the largest sub-segment, Data Warehouse, decreased by 0.9 percent to \$286 million in FY2017.



DoD Investment in Data Quality Helps Pave Way for Convergence of Big Data and AI

DoD is placing greater emphasis on data quality than it did in the past. A primary reason is that the Department is finding that data quality is oftentimes more important that data quantity. FY2015 marked a turning point when DoD began prioritizing data quality over data collection.

Annual spending on the two smallest sub-segments related to data quality, Data Architecture & Modeling and Distributed Processing Software increased the most over the last five years by 6.8 percent and 2.7 percent respectively.

System Integrators such as Northrop Grumman, Raytheon, Deloitte, Leidos and Booz Allen Hamilton benefited the most from DoD's increased spending on Data Architecture & Modeling. Northrop and Raytheon generated most of its business from Missile Defense Agency, while Deloitte performed work for Defense Health Agency (DHA). Leidos' largest customer was AFLCMC and Booz Allen Hamilton performed work mostly for Navy. For now, these companies are well positioned to drive the Third Offset Strategy forward by integrating big data technologies with decision making operating concepts.

One of those Big Data technologies is Distributed Processing Software. Insight Public Sector, the market leader, sells its products mostly to AFLCMC, Army, Navy and NETCOM. Integrators also sell Distributing Processing Software to DoD. Lockheed Martin, Booz Allen Hamilton and Northrop Grumman rank among the top ten sellers of Distributed Processing Software.

Data Architecture & Modeling



Annual Contract Obligations by Sub-Segment

Distributed Processing Software

Big Data Technologies Sub-Segments

Data Warehouse

Vendor Market Share

SAP Insight Public Boeing Raytheon NANA Dell Insight Public Sector Regional Secto Mythics Tetra Tech DLT Northrop Solutions CSRA Microsoft Genera Harris World Wide Product Fechnology Data Northrop Allen Technologies CACI RDM Leidos SRC SAIC Deloitte ManTech Leidos SK Telecom Teradata Raytheon CDW Harris

Exhibit 12: DoD prioritized data quality in recent years. Two closely related sub-segments, Data Architecture & Modeling and Distributed Processing Software had the largest spending increases since FY2012. Systems Integrators benefited the most from the spending increases, strengthening their positions for helping to integrate Big Data technologies and AI with agency operating concepts.



Analytics is the Largest Taxonomy Segment and Spending is Growing Modestly

With analytics it is possible to get answers from your data almost immediately -- an important facet of human and machine teaming. But what makes analytics different from traditional methods of analysis is the speed and efficiency it provides; AI is poised to augment the capability.

The convergence of AI and Big Data brings that speed and efficiency to entirely new levels, which is expected to completely alter existing methods of Business Analytics, Intelligence Exploitation and Data Analytics and related software. Until that happens, humans will team with machines to leverage advanced data science tools to squeeze the most out of Big Data and rely on analytics to surface the findings for decisive action.

DoD spending data trends shed light on this reality. Traditional methods of Intelligence Exploitation and Business Analytics are the largest sub-segments accounting for 72.7 percent of total segment spending since FY2012. One of those sub-segments, Business Analytics had the most spending growth of 7.9 percent over the last five years.

Govini has categorized Analytics into the following four sub-segments:

- Intelligence Exploitation data methods such as translating, evaluating and transforming raw intelligence data and information into useful forms
- Business Analytics data skills, technologies and practices used to gain insight
- Data Analytics examining large data sets in order to draw conclusions, increasingly with the aid of specialized systems and software
- Data Visualization Software software that abstracts data in schematic form and organizes for visual representation

Analytics Sub-Segments

Business Analytics	Data Analytics
Intelligence Exploitation	Data Visualization Software

FY17 Contract Obligations Compared to 5 YR CAGR by Sub-Segment



also had significant growth of 6.2 percent and 4.1 percent respectively.



Technical Engineering Contractors to Play Key Role in Integrating AI Big Data

DoD spending on Analytics has been relatively stable from year-to-year compared to other Taxonomy segments. While top-line spending dipped slightly from sequestration in FY2013, it recovered quickly to reach its highest level yet of \$2.4 billion in FY2017.

Service firms that have the ability to deliver technical data solutions have benefited from the stable spending, particularly those providing Intelligence Exploitation and Business Analytics. The Third Offset Strategy, however, calls for much of the mission work to be automated under the direction of a human operator.

AASKI Technology captured 15.6 percent of Intelligence Exploitation spending, the largest share of all contractors through its support of Army. BAE Systems follows as the close second capturing 13.4 percent of the market mostly from support of its Digital Electronic Warfare System (DEWS). While Processing, Exploitation and Dissemination (PED) services are likely to play a critical role in operating concepts for the foreseeable future, those firms that leverage cutting-edge technologies that induce automation rather than prop-up manual human analysis stand the most to gain. This human-machine teaming process will not only end-up delivering more effective solutions, it will also reveal where AI could be easily implemented.

Business Analytics is not much different than Intelligence Exploitation in that professional service firms that can gain access to broader swaths of data and organize them for real-time analytics and high-fidelity analysis will gain share in a market in transition. Leidos led capture of Business Analytics spending with 7.5 percent mostly from its work on Geospatial Research, Integration, Development, Support (GRIDS II) Program.

Analytics Sub-Segments

Intelligence Exploitation
 Business Analytics
 Data Analytics
 Data Visualization Software

Annual Contract Obligations by Sub-Segment



Vendor Market Share



two sub-segments have the most to gain from advancement in AI and Big Data.



DEPARTMENT OF DEFENSE CLOUD STANDARD MARKET TAXONOMY



Govini's *AI, Big Data and Cloud Taxonomy* present the Cloud category as the smallest by contract obligations accounting for only 16.7 percent of overall spending since FY2012. The category, however, is poised for rapid growth driven by the recent directive signed by Deputy Secretary of Defense Patrick Shanahan to accelerate enterprise Cloud adoption. Cloud and other forms of digital operating models are critical to storing, processing, managing and delivering the massive amount of data required for AI.

DOD Embarks on Transition to Cloud with Spending Up Across All Service Models

DoD has been slow to embrace Cloud technologies as a solution to their data challenges, but that is beginning to change. FY2016 marked a turning point for DoD Cloud, with Service Model spending having its sharpest rise on record by 31.2 percent to \$1.3 billion.

AI is among several factors prompting DoD to embark on its transition to Cloud. One use case showing great promise is the adoption of computer vision and machine learning technologies for concept search. The technologies allow users to search through troves of photos, images and other documents using visual components and concepts, instead of by file name or tag.

However, before AI can be applied at scale, DoD must transition to the Cloud and this is beginning to occur. Spending on the single largest Cloud Service Model, Infrastructure-as-a-Service (IaaS), increased the most since FY2012 by 9.8 percent to \$779.1 million in FY2017. Software-as-a-Service (SaaS) spending increased the second most by 8.3 percent to \$280.8 million and Platform-as-a-Service (PaaS) increased by 7.6 percent to \$347.7 million.

Govini has categorized Cloud Service Models into the following three sub-segments:

- Infrastructure-as-Service hosted infrastructure components traditionally present in on-premise data centers, including servers, storage and virtualization layer
- Platform-as-a-Service resilient and optimized environment on which users can install applications and data sets
- Software-as-a-Service hosted applications made available over the internet

Cloud Service Models Sub-Segments PaaS

laaS

SaaS

FY17 Contract Obligations Compared to 5 YR CAGR by Sub-Segment



Exhibit 15: IaaS, the largest Cloud Service Model sub-segment, had the most spending growth of 9.8 percent since FY2012. Spending on PaaS, the second largest sub-segment, increased by 7.6 percent and SaaS spending increased by 8.3 percent.

.govini

Cloud Spending Sees Sharp Rise in FY2016 and is Set to Continue its Strong Growth

DoD spending on Cloud is set to surge and government and industry are positioning for the investment funds to flow. Increased budget for Cloud and IT Modernization is one indication and a recently issued directive from the Deputy Secretary of Defense to accelerate enterprise Cloud adoption is another reason.

What makes Cloud providers uniquely positioned in the market is how they will leverage Learning & Intelligence technologies to make optimal use of data in an open environment while keeping the data secure.

This fact has prompted industry to make big bets in the form of mergers and acquisitions. DXC Technology, CSRA, Leidos and Booz Allen Hamilton have doubled down on their market position while others like Lockheed Martin and Harris strategically chose to put their chips elsewhere. Still others like General Dynamics, Northrop Grumman and CACI have yet to make moves that help them keep up with the evolving competitive landscape of combining forces.

Thus far, DXC Technology is the Cloud Service Model market leader, mostly from its position in IaaS. IBM, Leidos, Booz Allen Hamilton, General Dynamics CACI and SAIC rank among the top ten Cloud Service Model providers by revenue captured.

Commercial Cloud solutions sold through Value-Added Resellers (VARs) DLT Solutions, Carahsoft, Inforeliance and World Wide Technology present viable alternatives to on-premise private cloud networks managed by integrators.

Cloud Service Models Sub-Segments PaaS

laaS



SaaS

Annual Contract Obligations by Sub-Segment





IaaS accounted for 53 percent of total Cloud Service Model spending since FY2012 and PaaS accounted for 26.4 percent. Spending in the two sub-segment also grew steadily over the six years by 9.8 percent and 7.6 percent respectively.



23



Conclusion

The Third Offset Strategy aims to improve Joint Force battle network performance and restore a comfortable U.S. conventional overmatch against potential adversaries, thereby strengthening conventional deterrence. Toward this end, the Strategy outlines DoD's intent to pursue rapid advancements in the field of Artificial Intelligence and Autonomous Systems, in order to pave the way towards major advances in human-machine collaboration and combat teaming, if not a new military-technical revolution.

DoD investments in AI and autonomous systems have been steadily rising since FY2013, and they saw a boost in FY2016 and FY2017 after the formal announcement of the Third Offset Strategy in November 2014. By FY2017, Department spending on the three biggest associated technological categories—Artificial Intelligence, Big Data and Cloud—reached \$7.4 billion, which is 32.4 percent higher than the \$5.6 billion spent in FY2012. The strong growth reflects both new DoD technological applications as well as a number of successful technological applications available from the U.S. private sector in AI technologies. Exploiting AI-related research and development in the vibrant American commercial sector is a key aspect of the Third Offset Strategy and so is leveraging the national network of Federally Funded Research and Development Centers (FFRDCs) and University Affiliated Research Centers (UARCs).

In the Artificial Intelligence category, spending on AI Systems has gained traction within DoD, especially with regard to Virtual Reality for training and simulation and Computer Vision for ISR. Spending on Virtual Agents is also increasing. More advanced military AI Systems will rely heavily on research and spending in the Learning & Intelligence segment being conducted and funded primarily by Defense Research Agencies and Laboratories, notably DARPA. DARPA accounted for 28.5 percent of spending on the three most critical Learning & Intelligence sub-segments—Deep Learning, Machine Learning and Natural Language Processing.

Spending in these three sub-segments has demonstrated great promise in narrowly defined military AI Systems. However, there are constraints to moving towards more advanced applications. For example, more capable AI systems are almost entirely dependent on Advanced Computing, an area that China is vying for leadership with the U.S. Despite the increased DoD spending in Advanced Computing since FY2013, it seems evident DoD will have to increase its investments in Supercomputing, Neuromorphic Engineering and Quantum Computing just to stay ahead, much less lead, in this area.

The same goes for Big Data and Cloud, the other two big Third Offset categories. DoD must continue to make foundational investments in Big Data and Cloud to facilitate advances in machine learning, the key to realizing the full revolutionary potential of autonomous systems. DoD must continue its transition to Cloud and do better at leveraging Big Data technologies for collecting and processing data as well as make better use of data science for analytics. The recent directive signed by Deputy Secretary of Defense Patrick Shanahan to hasten DoD's transition to Cloud is thus a welcome development.

Unfortunately, DoD will not derive much benefit from its move to the Cloud without demanding much better data hygiene that results in stored data that has been processed, standardized and normalized for use. A concerted commitment to good data hygiene is the primary reason why



the U.S. private sector has achieved success in deploying AI applications. Leading AI companies like Google, Amazon, Facebook among others have access to endless data of the highest fidelity. In contrast, DoD struggles with data quality, data processing and data sharing. The Department's closed data architecture limits data sharing almost by design, which ultimately determines whether and how data can be cleansed, standardized and normalized. Ultimately this must change and Cloud and other forms of digital operating models provide the answer—both of which will require more investments.

However, on balance, a review of DoD spending on AI and autonomous system reveals a glass half full. DoD appears to be "putting its money where its mouth is" when it comes to pursuing Third Offset technologies. Hopefully, this report will help the new Under Secretary of Defense for Research and Development take stock of the Department's AI research and development portfolio and make the necessary changes to ensure continued progress in this important military arena.

These changes should be informed by the following question: are we spending enough on AI and autonomous systems, in the right areas, for the right outcomes? Both to improve its economy and its own military's importance, China recently released a national strategy to surpass the U.S. and become the world leader in AI theories, technologies and applications by 2030. Whether it wants it or not, the U.S. now finds itself in a major technological competition with a formidable rival. As Bob Work has asked, how will the U.S. respond to this "Sputnik moment"? Will the Chinese national plan be met with one of our own? Given the high stakes, one hopes so, as this competition will have very real consequences both in economic and military terms.

One thing is certain: Any national response must be driven by hard, accurate decision-grade information.

Methodology

Govini creates decision-grade information that allows clients to tackle their most difficult problems. Govini takes a unique taxonomic approach to breaking apart the market and its players, and provide insights only available through its Strategic Intelligence Platform. These analytic reports are designed to categorize Federal Government contract obligations and budgets into segments and sub-segments. Because some contracts are broad in scope, they may be included under multiple categories within a taxonomy to ensure an accurate, granular and evidence-based reflection of the market.



Govini is a big data and analytics firm committed to transforming the business of government through data science. The company's insights and analyses are utilized by Federal Contractors, Federal Agencies, Private Equity Firms and Hedge Funds to guide their strategies and uncover opportunities. Govini was founded in 2011 and has offices in Arlington, Virginia and San Francisco, California.

