SALVE REGINA UNIVERSITY

SENSING A CONTROL PROBLEM?

AUTONOMOUS UNMANNED COMBAT AIR SYSTEMS AND HUMAN CONTROL

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE HUMANITIES PROGRAM

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

SHAWN A. KALIS

NEWPORT, RHODE ISLAND

MAY 2016

SALVE REGINA UNIVERSITY

GRADUATE STUDIES

This dissertation of Shawn A. Kalis, entitled "Sensing a Control Problem? Autonomous Unmanned Combat Air Systems and Human Control," submitted to the Ph.D. Program in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Salve Regina University, has been read and approved by the following individuals:

Reader: ______________________ May 9, 2016
Michael A. Budd, Ph.D. (Date)

Reader: ______________________ May 9, 2016
Troy T. Catterson, Ph.D. (Date)

Mentor: ______________________ May 9, 2016
Craig Condella, Ph.D. (Date)

Director,
Ph.D. Program: ______________________ May 9, 2016
Michael A. Budd, Ph.D. (Date)

Vice President
for Academic
Affairs: ______________________ 5/9/16
Scott Zeman, Ph.D. (Date)

TABLE OF CONTENTS

DEDICATION

To my wife Dawn and the Sensational Seven,
for the constant motivation to "keep walking!"

In memory of Captain Eric Shaw, Ph.D., USCG
1957-2016
My advisor, teacher, mentor, and brother in arms.

ABSTRACT

Military unmanned aerial system technology, remote sensing systems, and artificial intelligence all play critical roles in creating the possibility for completely autonomous unmanned combat air systems. The integration of these technologies is significantly affecting the human-technology relationship on the battlefield. Not only are these technologies decreasing the level of direct human interaction in warfare, they are also significantly increasing the physical distance between humans and machines. This increased distance raises questions regarding the ability of humanity to maintain control over autonomous weapons systems. The development of these autonomous technologies appears to validate the position of traditional technological determinists such as Martin Heidegger and Jacques Ellul, as well as contemporary thinkers such as Raymond Kurzweil, Stephen Hawking, and Elon Musk, who propose that artificial intelligence technology is not a thing that can be controlled, but rather an influence or force that transcends human ability to maintain control over, and that any sense of control is merely an illusion. This work defends the thesis that while integrated technologies are significantly contributing to the evolution of fully autonomous unmanned combat air systems, the existence and application of these technologies do not by default support the hard technological deterministic view that technology is a force or influence beyond human control, nor does the evolution of such technologies necessarily threaten humanity's ability to maintain control.

Chapter One: Introduction

"The emotional and moral disengagement of the cubicle warrior may increase in the future, due to a noticeable shift from controlling to monitoring."

— Lamber Royakkers and Rinie van Est

This dissertation defends the thesis that while the integration or blending of unmanned aircraft system technology, remote sensing systems, and artificial intelligence (AI) play a critical role in the evolution of the U.S. military's autonomous unmanned combat air system (UCAS), the existence and application of these technologies do not necessarily support a technological deterministic view that humanity's ability to maintain control is threatened or transcended. Instead, the findings, with regards to the particular issue of the autonomous UCAS, will reveal that they run counter to the traditional hard technological determinist claims that technology is moving beyond human control.

Technology evolving beyond human control and threatening to convert humans from active participants to passive recipients has been a concern for humans since the Luddites of the industrial revolution. The theme of technology taking control is popular in many genres of literature and film. In modern times, Stanley Kubrick's famous 1968 film *2001: A Space Odyssey*, inspired by Arthur C. Clarke's 1948 short story, "The Sentinel," poignantly introduced the topic of technology's challenge to human control with the HAL 9000, a heuristically programmed algorithmic computer that has artificial intelligence and acts as the antagonist in the story. In D.J. Caruso's 2008 film *Eagle Eye,* the Autonomous Reconnaissance Intelligence Integration Analyst (ARIAA) super-computer determines that humans are no longer capable of making correct decisions; therefore, it breaks away from human control to embark on a mission to protect the spirit

of the Declaration of Independence by attempting to eliminate the executive branch of the U.S. government. One particular science fiction film extremely relevant to this study—because it depicts the struggle for control between humanity and technology that appears to be coming to fruition today—is Rob Cohen's 2005 film *Stealth*. The movie focuses on a runaway autonomous unmanned combat air vehicle (UCAV) called "EDI," short for the Extreme Deep Invader. EDI contains multiple remote sensing systems and is completely controlled by an artificial intelligence system with quantum processing and a neural network. The UCAS has the capability to outperform human pilots and, in the course of the film, develops self-awareness and refuses to obey human commands, becoming an adversary to human pilots. Film director Cohen, reflecting on the technology-humanity relationship, stated that he began to realize that "within the story, it contained one of the great issues of our time . . . that as technology develops to the point of independence, who is in control and what are the possibilities, and especially as applied to war . . . ?" (Fischer 2005).

On May 14, 2013, off the coast of Virginia, the Northrop Grumman Corporation and the crew of the aircraft carrier *USS George H.W. Bush* (CVN-77) brought Cohen's science fiction film one step closer to reality. On that date the Northrop Grumman X-47B Unmanned Combat Air System Demonstrator successfully completed its first autonomous catapult launch as part of the U.S. Navy's UCAS Carrier Demonstration (UCAS-D) program (Vergakis, 2013). Three days later this event was followed by the first carrier-based touch-and-go landings by an autonomous UAV, and on April 22, 2015, the X-47B conducted the first ever autonomous aerial refueling (Northrop Grumman

2015, 1). What was once purely science fiction appears to be developing into reality, and the questions still remain: “Who is in control?” and, “What are the possibilities?”

What Does It Mean to Be Human in the Age of Technology?

“It has become appallingly obvious that our technology has exceeded our humanity.”
— Albert Einstein

The motivation for this examination is the desire to discover the impact that a specific human-technology relationship has on society and culture, and to apply the findings to help provide one of many possible answers to the question “What does it mean to be human in an age of technology?” This inquiry focuses on the element of control in the human-technology relationship, and the impact that the issue of control has on this relationship. Control has been a consistent theme in the evolution of humanity, as evidenced in the history of mankind’s struggle to control one another, the elements, the environment, and even one’s self. Examining the issue of control as it relates to the human-technology relationship is vital to understanding what it means to be human in an age of technology.

With this broad question in mind, this study will examine the technologies used in developing autonomous aerial weapon systems, like the X-47B UCAS, in order to address the human-technology control issue. Through an examination of the human-technology interface in the context of 21st-century warfare and unmanned weapon systems development, specifically the UCAS, this dissertation helps determine the impact that this evolving technology has on the human operator’s ability to maintain control of the technology. Using the U.S. military’s development of the UCAS as a case study, this

work specifically examines the impact of the integrated technologies of the unmanned aircraft system, the remote sensing system, and AI technologies on the evolution of the UCAS. The impact of these integrated technologies on the human operator is examined to determine the level of influence on the operator's ability to maintain control of the technology. The specific question addressed is: *"Do these integrated technologies, in the specific form of a UCAS, support a technological determinist argument that humans are losing control, or have no control, over the technology?"*

There is a broad spectrum of viewpoints regarding the relationship between humanity and technology in the realm of control. One end of the spectrum asserts that technology is and always will be under the control of human beings and society; this falls within the broader philosophical category of social determinism. This position is held by philosophers such as Don Ihde, Jurgen Habermas, and Leila Green. The opposite end of the spectrum affirms that humans never had, and never will have, the ability to control technology; this position lies within the realm of the philosophical position of technological determinism. Philosophers such as Martin Heidegger, Jacques Ellul, and Raymond Kurzweil adhere to this view. Between both extremes lies a wide variety of opinions and positions on the influences of technology on humanity, and vice versa. Philosophers such as Langdon Winner, John McDermott, and Robert Heilbroner espouse varying views of the different degrees that technology and humanity influence each other. (This spectrum of viewpoints on the relationship between humanity, technology, and control will be addressed in greater detail in chapter two.) At the heart of the issue between the various human-technology relationship spheres is the question of control.

This study examines the hard technological deterministic position that the evolution of autonomous technology is an example of a force or influence that is beyond humanity's ability to control. *The position advanced in this dissertation is from a social deterministic view, that technology, while becoming more complex and autonomous, is a human creation, a tool that remains under human control.* While it is true that the unmanned aircraft, remote sensing, and AI integrated technologies are making indispensable contributions to the creation of autonomous UCASs, they do not appear to support the technological determinist's view that autonomous UCAS technology is actually evolving beyond human control.

The findings of this dissertation may be directly applied to discussions and initiatives addressing current and future military, political, and social implications of the use of autonomous unmanned weapon systems in modern warfare. The findings not only apply to current moral questions concerning the control of autonomous military technology, but also apply to U.S. policy questions concerning the use of such systems in the future. In addition, the findings of this dissertation may also be used by U.S. senior-level policy and military decision-makers who are directly involved with developing autonomous military technologies, providing information that can encourage them to keep the humanity-technology relationship in consideration in order "to work for a world that is harmonious, just, and merciful."[1]

[1] In addition to encouraging students to seek wisdom and promote universal justice, the Salve Regina University mission statement encourages students "to work for a world that is harmonious, just, and merciful" (Salve Regina 2016).

Key Terms Defined

"Nothing has meaning without context."

— Joe Aldrich

In order to examine the issue of integrated unmanned aircraft, remote sensing, and AI technologies and their influence on humanity's ability to maintain control, a few key terms must be defined: *technology*, *unmanned aircraft system*, *remote sensing*, *artificial intelligence*, and *autonomy*. While there are many technical terms used throughout the course of this study, some of these terms are more problematic than others in that there is less unified agreement regarding their definition.

Technology

According to the Oxford English Dictionary, the word *technology* originates from the Greek τεχνολογία (technology), a combination of two words: τέχνη (téchnē) and -λογία (-logia). The meaning of τέχνη is the subject of much discourse and debate in the field of the philosophy of technology; however, in the context of the English vernacular usage of technology, most dictionaries and encyclopedias define τέχνη (téchnē) as art, skill, craft, or the way in which something is gained.[2] The word -λογία (-logía), is less debated in the field, and has the accepted meaning of "study of." Like many terms used and debated among scientists and philosophers, the word *technology* is defined in different ways. Cultural theorist Paul Virilio describes technology as a mystery in *Pure War*, stating that "any examination of technology immediately gives rise to

[2] The more complex definitions of τέχνη as articulated by philosophers of technology such as Martin Heidegger and Jacques Ellul, and the associated issues related to technological autonomy, will be addressed in later chapters.

misunderstandings" (Virilio 2008, 37). Stanford Professor Stephen J. Kline stated that, "In the late 20th century, there is only one thing most people agree about concerning technology—it is important" (Kline 1985, 215). However, despite the ambiguities and misunderstandings, there are enough commonalities that surface among theorists and practitioners that allow for a working definition sufficient for the purposes of this study. Economist John Kenneth Galbraith includes systemization and science in his definition in *The New Industrial State,* defining technology as "the systematic application of scientific or other organized knowledge to practical tasks" (Galbraith 2007, 14). Technology philosopher Carl Mitcham, in his book *Thinking Through Technology: The Path Between Engineering and Technology*, gives the simple definition of technology as "making and using artifacts," and then throughout the rest of his treatise demonstrates the "narrow and broad senses" of the term as used by engineers and humanities scholars (Mitcham 1994, 1-15). Philosophers Mary Tiles and Hans Oberdiek, in *Living in a Technological Culture*, include "ways of doing and making which are both affected by and affect ways of thinking" in their definition; they insist that artifacts, techniques, roles, and practical knowledge be included in the definition (Tiles 1995, 10). In *Science and Technology in World History*, History of Science professors James McClellan and Harold Dorn underscore the practical aspect of technology but separate it from science, which they state is "more theoretical and philosophical" (McClellan 2006, 47-54). Science and Technology philosopher Don Ihde emphatically states that "definitions are not neutral," and therefore gives technology a "middle-sized" definition, insisting that the definition contain three components:

> First, we shall insist that a technology must have some concrete component, some material element, to count as technology. And, second, a technology must enter into some set of praxes—'uses'—which humans may make of these components. And third, we shall take as part of the definition, a relation between the technologies and the humans who use, design, make, or modify the technologies in question. (Ihde 1993, 47)

The important point of Ihde's definition is that in addition to the application of knowledge in creating an artifact for practical purposes, there is a relationship between humanity and technology. Many definitions include Ihde's components, emphasizing the human-technology interaction in the application of science or knowledge to create a particular capability for a specific, usually practical, purpose. Therefore, assuming that there will be a relationship of some kind between the human and the technology created, the definition of technology used for this study is as follows: *Technology is the application of knowledge that produces an artifact or provides a capability for practical purposes.* The different philosophical points of view taken on the relationship between technology and humanity, especially as it is applied to the aspect of technological autonomy and control, will be discussed in more detail later in the study.

Unmanned Aircraft System

Among the many different types of unmanned military technologies, the pilotless aircraft has played a significant role, primarily because of its successes in the U.S. Global War on Terrorism and Overseas Contingency Operations in Iraq and Afghanistan. The official term used by the Department of Defense when referring to a craft in the air without a pilot onboard is unmanned aircraft (UA), which it defines as "an aircraft that does not carry a human operator and is capable of flight with or without human remote control" (U.S. Department of Defense Joint Publication (JP) 1-02 2010, 252). All of the

associated technology needed to operate a UA is termed an Unmanned Aircraft System (UAS), which is officially defined as "that system whose components include the necessary equipment, network, and personnel to control an unmanned aircraft" (U.S. Department of Defense JP 1-02 2010, 252). A UAS that is used as a military weapon system is called an unmanned combat air system (DARPA 2004, 1; U.S. Navy 2014, 1). Other terms often used by the DOD, other services, organizations, and agencies to refer to UA and UAS technology include: drone, optionally piloted aircraft, remotely operated aircraft, remotely piloted aircraft, remotely piloted vehicle, unmanned aerial system, unmanned aerial vehicle, and unmanned combat aerial vehicle.[3] This study will predominately use the terms UA, UAS, and UCAS; however, some of the other terms may be used depending on source references and context.

Remote Sensing System

Remote sensing is a critical component of unmanned and autonomous systems operations because remote sensors operate as eyes to the external world for machines.[4] The most commonly accepted definition of remote sensing among practitioners in the field is the action of observing something from a distance, most often the earth, usually

[3] Many of these terms substitute the use of "aircraft" with the word "aerial." While the only terms beginning with "unmanned" in the Joint Publication 1-02, *Department of Defense Dictionary of Military and Associated Terms* (8 November 2010, As Amended Through 15 January 2016) are for the Unmanned Aircraft and Unmanned Aircraft System, the publication itself uses the term "unmanned aerial vehicle (UAV)" elsewhere in the document (see page 107 under the definition for imagery or the section entitled "Abbreviations and Acronyms," where UAV is defined as "unmanned aerial vehicle" numerous times). The earlier 12 April 2001 (As Amended through 31 October 2009) edition of JP 1-02 contained definitions for "remotely piloted vehicle" and "unmanned aerial vehicle," but these have been removed from the 2010 edition. This illustrates that many of the earlier terms previously used are still popular and are used to describe airborne craft without a pilot physically onboard.

[4] The term "remote sensing" was first used in the U.S. by Evelyn Pruitt of the U.S. Office of Naval Research in the 1950s and "is now commonly used to describe the science—and art—of identifying, observing, and measuring an object without coming into direct contact with it" (Graham 1999, 1).

from the air or space, using special instrumentation that measures reflected energy (e.g., heat from the sun radiating off an object) or emitted energy (e.g., heat from a car engine). The U.S. Geological Survey defines remote sensing as the "process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation at a distance from the targeted area (U.S. Geological Survey 2012). The National Geospatial-Intelligence Agency (NGA), the U.S. Department of Defense's National Functional Manager of imagery and geospatial intelligence (GEOINT), describes remote sensing as:

> . . . acquiring information about an object or target using a device that is *not in physical proximity with the object under study*. Remote sensing occurs through an interaction between some form of electromagnetic energy (natural or man-made) and objects and phenomena on or above the earth or below the earth's surface. Remote sensors can employ a number of technologies including electronic, optical, electro-optical, chemical, or mechanical systems, either individually or in combination. The information can be recorded or analyzed in either imagery or non-imagery formats using digital or analog means. (U.S. National Geospatial-Intelligence Agency 2006, 19; emphasis added)

Remote sensing information is categorized into two types: literal and nonliteral. Literal information is classified as visible imagery, e.g., a black and white (panchromatic) or color image that gives the viewer a literal or actual depiction of the object or scene (U.S. NGA 2006, 19). Nonliteral remote sensing consists of systems that collect data using sensors that produce information that cannot be interpreted by the human eye and cognitive systems, i.e., it is not recognizable as a literal depiction of the object or scene (U.S. NGA 2011, 17). Examples of nonliteral remote sensing include synthetic aperture radar (SAR), infrared (IR), and hyperspectral imagery (HSI), techniques that will be described in detail later in this study.

The definition of remote sensing used for this study is as follows: *the process of acquiring literal and nonliteral information about an object or target on, above, or below the earth or water using sensors onboard an unmanned aircraft (UA).* The remote sensing hardware, software, and aerial sensor technologies used to collect data; the data transmission and communication hardware and software used to transmit and receive commands and remote sensing data to and from the sensors onboard a UA; the hardware and software used to process, analyze, exploit, store, and disseminate remote sensing data; as well as all of the associated infrastructure and personnel required to operate the sensors and manipulate the data, together comprise the remote sensing system.

Artificial Intelligence

Within the computer science community there is little debate that English mathematician Alan Turing was one of the first persons to focus on the computer programming aspects of artificial intelligence (AI). Work on creating intelligent machines began in earnest in the post-World Word II era, and Turing gave a lecture to the London Mathematical Society in 1947 that included the subject of machines and intelligence (Turing 1986, 106). Soon after, computer programming research in the field of AI significantly increased (McCarthy 1998, 3).[5]

[5] In contemplating artificial intelligence, Turing addressed the concept of whether machines could think in an article in the journal *Mind* in 1950. In the article, he addresses the question not by defining the terms "machine" or "think," but by re-describing the problem using a modified version of a popular three-person guessing game—replacing one player with a computer—that would test the machine's ability to display intelligence that would be indistinguishable from that of a human player. His game became known as the "Turing Test." (Turing 1950, 433, 434) The Turing Test has been at the heart of many philosophical discussions about artificial intelligence and can be seen referenced in many computer science works on the topic as well. A good example of the types of discussions generated can be seen in NYU cognitive science professor Gary Marcus's *The New Yorker* article titled "What Comes After the Turing Test?" (Marcus 2014).

Like technology, the term "artificial intelligence" poses some significant challenges, especially with the word *intelligence* as a component. The word *artificial* is less problematic as it is usually accepted to mean man-made as opposed to occurring naturally. However, one of the most significant issues with determining an acceptable working definition of artificial intelligence is settling upon an acceptable definition of human intelligence. Cognitive scientist Marvin Minsky, pioneer researcher in the field of AI and founder of the MIT AI Laboratory, readily admits that there is "no generally accepted theory of 'intelligence'" (Minsky 1960, 4), and futurist Ray Kurzweil underscores this statement by reaffirming that "there appears to be no simple definition of intelligence that is satisfactory to most observers" (Kurzweil 1992, 18). Because this study focuses on computerization of human intelligence in relation to autonomous weapon systems and control, the definition for intelligence developed by computer scientists John Laird, Allen Newell, and Paul Rosenbloom, early pioneers in a Defense Department AI research project called SOAR, is important to consider.[6] The goal of the SOAR program was to develop an AI software system capable of solving problems and learning in ways similar to human beings. The team stated that the goal of its efforts in working to create a cognitive architecture was to "provide the foundation for a system capable of general intelligence behavior," the characteristics for which they described as the ability to:

> (1) Work on the full range of tasks, from highly routine to extremely difficult open-ended problems, (2) employ the full range of problem solving methods and

---

[6] "SOAR: An Architecture for General Intelligence" was the 1986-1991 artificial intelligence and psychology research effort supported by the Computer Sciences Division and the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research; the Defense Advanced Research Projects Agency (DARPA); Air Force Avionics Laboratory; National Institutes of Health; and the Sloan Foundation. The specific work cited here is from Technical Report AIP-9 (Laird 1987).

> representations required for these tasks, and (3) learn about all aspects of the tasks and its performance on them. (Laird 1987, 1)

This early definition of intelligence is closely aligned with one agreed on by 52 experts in intelligence and allied fields and published in a 1994 *Wall Street Journal* statement entitled "Mainstream Science on Intelligence."[7] The definition agreed on in that statement, and the one which will be used for this study, is as follows:

> Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings—"catching on," "making sense" of things, or "figuring out" what to do. (Statement, 1994)

With this definition of intelligence as the baseline, a working definition for artificial intelligence suitable for this study can be constructed.

It is also generally accepted by members of the AI community that computer and cognitive scientist John McCarthy first coined the term artificial intelligence when organizing the 1956 Dartmouth Summer Research Conference on Artificial Intelligence. However, despite the early use of the term and subsequent progress made in the field, there is still no commonly accepted definition for artificial intelligence (Murphy 2000, 15). John McCarthy defines it as "the science and engineering of making intelligent machines, especially intelligent computer programs" (McCarthy 1998, 2). Minsky also sees AI as making machines do things that require human intelligence (Minsky 1968, 23), and Kurzweil defines it in a similar manner, although he emphasizes the controversial

[7] The "Mainstream Science on Intelligence" article was written by professor of psychology Linda Gottfredson and signed by her and 51 other university professors specializing in intelligence (Statement, 1994).

nature of trying to define AI due to the difficulties of defining intelligence. British psychologist Michael Eysenck underscores the complexities of this human-technology relationship with his definition of AI by emphasizing that this is an activity that attempts "to develop complex computer programs that will be capable of performing difficult cognitive tasks" (Eysenck 1990, 22).

As with technology, there are many theories associated with AI, and like technology, there are enough commonalities to allow us to form a working definition for this study. Again, assuming that there will be a relationship of some kind between the human and the technology created, the definition of artificial intelligence used for this study is: *Artificial intelligence is the creation of computer software and hardware that allows a machine to perform functions that would require intelligence when performed by a human being.* The degree to which a machine can do this without human intervention is addressed in the following definition of autonomy.

Autonomy

The root of the word *autonomy* is derived from ancient Greek and means "having its own laws," from the words *autos*-, meaning "self," and *nomos*, meaning "law."[8] As with the definitions of technology and AI, autonomy is problematic in that there is no

[8] Autonomy for this study is initially approached ontically in order to provide a working definition that focuses on autonomous weapon systems as technologies that are real, currently exist, and are used in the United States Department of Defense. This definition focuses on the concept of self-governance as applied to technological autonomy and the ability of a machine to use artificial intelligence in order to operate independently from human control. The ontological arguments associated with the development and evolution of autonomous weaponry is addressed in the chapter on technological determinism. In that section and elsewhere, the moral philosophical elements offered in the Kantian sense, i.e., the values and desires associated with an agent's capacity to impose the moral law on oneself, will be discussed; however, a comprehensive study of the arguments related to technological autonomy in relation to a moral law are beyond the scope of this study.

agreement when applied to technological systems. Dictionary definitions of autonomy underscore self-government, independence, and self-determination, especially in terms of existence, e.g., being independent or self-governing. These are also key elements when defining machine autonomy. Engineer and computer scientist Robin Murphy defines technological or robotic autonomy as a system that "can adapt to changes in its environment (the lights get turned off) or itself (a part breaks) and continue to reach its goal" (Murphy 2000, 4).

The Department of Defense (DOD) also includes the capability and freedom to self-direct and make choices in its definition, although the DOD does emphasize that the human-technology relationship is "defined by policy and operational requirements" (Stone 2011, 3). The DOD Unmanned Systems Integrated Roadmap FY2011-2036 draws a clear distinction between an automatic and an autonomous system and the level of human interaction involved with each. The distinction is described here as follows:

> Automatic systems are fully preprogrammed and act repeatedly and independently of external influence or control. An automatic system can be described as self-steering or self-regulating and is able to follow an externally given path while compensating for small deviations caused by external disturbances. However, the *automatic system is not able to define the path according to some given goal or to choose the goal dictating its path*.
>
> By contrast, autonomous systems are *self-directed toward a goal* in that they *do not require outside control*, but rather are g*overned by laws and strategies that direct their behavior*. Initially, these control algorithms are created and tested by teams of human operators and software developers. However, if machine learning is utilized, autonomous systems *can develop modified strategies for themselves by which they select their behavior*. An autonomous system is *self-directed by choosing the behavior it follows* to reach a human-directed goal. (Unmanned Systems Integrated Roadmap FY2011-2036, 43; emphasis added)

DOD Directive 3000.09, *Autonomy in Weapon Systems,* also makes this distinction, defining an autonomous weapon system as: "A weapon system that, once

activated, can select and engage targets without further intervention by a human operator. This includes human-supervised autonomous weapon systems that are designed to allow human operators to override operation of the weapon system, but can select and engage targets without further human input after activation" (U.S. Department of Defense Directive 3000.09 2012, 13). The FY2013-2038 Unmanned Systems Integrated Roadmap echoes these definitions, also stating that automatic systems require human control, and autonomous systems are able to make decisions and react without human interaction (U.S. Department of Defense Unmanned Systems 2013, 68).

This distinction given by the DOD is important, first because it differentiates between automatic "fire and forget" weapon systems, such as cruise missiles or guided munitions, and unmanned or autonomous weapons systems, such as the X-47B UCAS. Second, because it clearly states that an autonomous system is *self-governing* in the context of its operating environment, i.e., it does not require human intervention for processing and responding to the environment and developing *modified strategies* to achieve its human-directed goal. This definition of autonomy exhibits the elements of AI previously given by the SOAR team: the ability to (1) work on a range of tasks, (2) employ problem-solving methods, and (3) learn about all aspects of the tasks and its performance on them. While both automatic and autonomous systems operate independently of external influence or control, only the autonomous system applies AI in order to adapt to its environment without human interaction.

Like the different degrees of intelligence, there are different degrees of autonomy, usually described in levels or scales that are based on the amount of human interaction. As there is no agreement on the definition of autonomy, the levels vary from organization