Artificial Intelligence and Administrative Discretion: Implications for Public Administration

Thomas J. Barth and Eddy Arnold

The American Review of Public Administration 1999; 29; 332
DOI: 10.1177/02750749922064463

ARTIFICIAL INTELLIGENCE AND ADMINISTRATIVE DISCRETION Implications for Public Administration

THOMAS J. BARTH
University of North Carolina at Wilmington

EDDY ARNOLD
University of Memphis

Advances in the field of artificial intelligence (AI) are leading to a new level of computing in which systems will have the capability to act as autonomous agents and learn to learn independently, assess their environment, and think with values, motives, and emotions. Reflection on the dialogue in the AI literature and implications for public administration raises issues concerning a number of classic dilemmas relevant to administrative discretion, including responsiveness, judgement, and accountability. After a brief overview of the AI field to provide context, this article addresses each of these themes in turn and concludes with a summary discussion on the potential benefits and dangers of AI for the field of public administration.

> It is not my aim to surprise or shock you. . . . But the simplest way I can summarize is to say that there are now in the world machines that think, that learn and that create. Moreover, their ability to do these things is going to increase rapidly until—in a visible future—the range of problems they can handle will be coextensive with the range to which the human mind has been applied.
>
> —*Simon in Dreyfus (1994, p. 81)*

It can be instructive to take a fresh look at enduring issues in the field of public administration by using the perspective of other disciplines. One such issue is independent decision making by public administrators, which has always been a contentious point for governance theorists. Although some degree of administrative discretion to execute law is inescapable in the modern administrative state, the question of how much is appropriate or legitimate without subverting legislative authority and accountability is a classic question. For example, Cook (1995) notes James Q. Wilson's argument that increasing bureaucratic discretion is necessary to contend with the vague, polycentric problems pursued by an activist democracy. In contrast, Theodore Lowi (1993) has warned that increasing levels of discretion is a certain formula for patronage government and unaccountable bureaucracies beyond the rule of law.

AMERICAN REVIEW OF PUBLIC ADMINISTRATION, Vol. 29 No. 4, December 1999 332-351

However, these theoretical debates should not be considered in a vacuum. As technology continues to produce increasingly sophisticated tools that enhance decision-making capabilities, it is important to continually revisit such arguments and consider potential implications of emerging technology.

Indeed, this article was prompted by a graduate student in a masters of public administration (MPA) program who is a professional communications engineer. In reaction to readings from an administrative ethics class, he noted that with the emergence of artificial intelligence (AI) technology, administrative discretion arguments will become moot. He was referring to the development of new levels of computers, which AI experts claim will have the capability to act as autonomous agents that can learn to learn independently, assess their environment, and think with values, motives, and emotions.

However implausible such machine intelligence may seem today, reflection on the dialogue and developments in the field of AI focuses attention on a number of dilemmas relevant to administrative discretion, such as responsiveness, judgement, and accountability. After a brief overview of the AI field to provide context, the article addresses each of these themes in turn and concludes with a summary discussion on the potential benefits and dangers of AI for the field of public administration.

THE LENS OF AI

As the term suggests, AI refers to the pursuit of machine or computer intelligence that approximates the capabilities of the human brain—or the "science of making machines do things that would require intelligence if done by men" (Marvin Minsky in Yazdani & Narayanan, 1984). The essence of AI is captured by the following statement: "Today, machines solve problems mainly according to the principles we build into them. Before long, we may learn how to set them to work upon the very special problem of improving their own capacity to solve problems" (Minsky in Dreyfus, 1994, p. 81).

As with any emerging technology, there is much debate over the potential capabilities and limitations of AI. Dreyfus (1994), one of the most cited critics of AI, argues that any artificial entity will ultimately fall short of human intelligence because it can never ultimately simulate the human experience. He notes that "computers can only deal with facts, but man—the source of facts—is not a fact or set of facts, but a being who creates himself and the world of facts in the process of living in the world" (pp. 290-291). Trefil (1997) argues that machines will never approach the capabilities of the human brain because of fundamental, functional differences between the cerebral cortex and digital computers. Others see this view as unnecessarily pessimistic and cite the undeniable progress toward artificial systems that are, in effect, autonomous agents—that is, machines that can independently think, learn, and solve problems (Franklin, 1995; Kurzweil, 1990). Indeed, Suplee (1997) concludes, "After a decade of

setbacks AI enthusiasts are again growing optimistic. Though they are still a long way from replicating the complexity of the human psyche, few theorists now claim that machine intelligence is impossible" (p. 85). Furthermore, these systems also may have the ability to incorporate human qualities such as values, motives, and goals (Crevier, 1993; Gill, 1986; Kurzweil, 1999; Waldrop, 1987, Yazdani & Narayanan, 1984). Even in this arena, AI critics do not question the technical feasibility of such machines but rather the morality of such endeavors (Sack, 1997).

To put the evolution of AI in context, it is useful to think of three levels of computer capability. We understand the use of computers as decision-making support, where they are useful as sophisticated, high-speed calculators. This first level of computing is what one typically sees in organizations today, often in the financial management arena. We are also becoming more familiar with the use of computers as expert systems, which we can think of as the second level of computing, where computers are able to apply sophisticated rule-based systems. An example of an expert system in the popular press today is Deep Blue, the IBM system that recently defeated world chess champion Gary Kasparov. A sampling of current applications in several fields provides a sense of this second level of expert systems.

In the field of medicine, expert systems such as MYCIN are being developed with the ability to quickly absorb information about particular patient cases (e.g., history, symptoms, laboratory test results) to help physicians formulate hypotheses, build evidence, and ask questions (Winston, 1992). In the military, systems such as the Integrated Defense System (IDS) are being designed to determine optimum response times, initiate appropriate countermeasures, and simulate both friendly and enemy aircraft (Daily Defense News Capsules, 1995; Hunter, 1995; Miller, 1994). In law enforcement, software called Brainmaker is being used as an automated warning system that uses an expert system to identify police officers who display patterns that could lead to crime and corruption (Seibel, 1994). A final example is in mental health, where an expert system called the Good Mood Program is being used to help users talk though personal problems they are experiencing (Ansley, 1994).

AI: A New Level of Computing

At both of these first two levels, the computers are only providing information based on what they are fed, although they are applying sophisticated logic and making connections independently. Although impressive, even the second level of expert systems is not the ultimate goal of the AI field. Franklin's concept of an autonomous agent with values and motives is another level entirely, because it means machines that have the capability to go beyond initial programming and learn to learn. For example, Artilects is a company devoted to the commercial development of next-generation AI software and holds

exclusive license to advanced AI technology called OSCAR. At present, OSCAR is being used to support medical decisions and credit analysis, but the company states,

> We plan to have OSCAR learning for himself within the next five years, using his current skills at defeasible reasoning as a basis for an inductive system of machine learning whereby OSCAR will create his own knowledge from raw data, and then learn as he goes along. (World Wide Web)

Furthermore, implicit in the ability to learn independently is the ability to exercise judgment as the situation changes. Waldrop (1987) notes that even the most sophisticated expert systems lack any semblance of common sense. Returning to the example of the Mycin expert system, he explains,

> Mycin knew a lot about diagnosing diseases. But it understood nothing whatsoever about anatomy or physiology and could not conceive that its patient was a human being. One of its rules told it to avoid giving tetracycline to patients under eight years of age. So it never did. But Mycin had no way of knowing why the rule was valid. (Tetracycline can produce dental staining during tooth development and may depress bone growth.) So Mycin couldn't explain the rule to the user or know when to break the rule (when, for instance, the child's life was so threatened that the cosmetic side effects were negligible). (p. 44)

Recognizing this limitation of expert systems, AI researchers are exploring how humans develop common sense through the use of analogy, trial and error, and reasoning by default. Grupe, von Sadovsky, and Owrang O. (1995) describe the concept of fuzzy logic and the development of fuzzy systems, which allow AI technology to adjust to changing conditions and excel in combining partially accurate, qualitative assessments of situations into decisions. Such developments are producing discussions about the hypothetical advantages and disadvantages of AI in areas previously considered much too subjective and sensitive for computers, such as the use of artificially intelligent judges and mediators in resolving socioscientific disputes in the field of law (Raghupathi, 1991; Spagnoletti, 1987), replacing astronauts with robots (Southerst, 1992), grading academic essays (Pappas, 1998), managing prison sentencing processes (Peterson, 1993), and providing investment advice (Pech, 1999).

Further examples of AI research and applications will be discussed, but the purpose of this article is not to analyze the enormous literature on AI or assess the debate over the scientific limitations of AI. Rather, the following discussion assumes progress toward this third level of computing and examines the implications for the field of public administration, particularly on the exercise of administrative discretion. Such reflection represents more than just an academic mental exercise; recent history demonstrates that the full social and political implications of new technology are rarely considered until after the fact, producing great societal stress (e.g., atomic energy, medical technology prolonging

life, genetic engineering). Furthermore, just as AI research is encouraging scholars in related fields such as philosophy and psychology to examine anew how humans think, reason, and learn, such inquiry also provides incentives for scholars in public administration to examine anew how we govern, make decisions, and provide services to the public.

From the perspective of administrative discretion, the potential significance of this third level of computing can be understood under three themes. First, the ability to program machines with values and motives suggests the potential to improve the rationality of decisions through tools that can apply a known or specified ranges of values or biases (theme of responsiveness). Second, the ability to develop machines that can sense subtle aspects or changes in the environment suggests tools that can make political or situational assessments (theme of judgment). Finally, machines that can learn to learn independently suggest a tool without precedence that may exceed the capacity of humans to scan the environment, assess situations, and make decisions in a timely manner without human supervision (theme of accountability). Reflection on each of these themes through the lens of AI reveals a number of questions, benefits, and dangers that should not be ignored as decision-making technology continues to emerge at an often unpredictable pace.

RESPONSIVENESS

One of the major dilemmas with public administrators exercising discretion is that the public must rely on the "public interestedness" of the administrator when we know that people and agencies can be motivated by self-interest. Indeed, Lowi (1993) notes, "The assumption of selfish interests is probably the only thing on which all political scientists agree" (p. 262). Under the leadership of James Madison, the Founders therefore created a government of checks and balances as well as popular elections to address self-interested behavior. The concern is that unelected administrators with discretion are beyond the reach of these checks, so they must be carefully constrained. Furthermore, even if they are not motivated by conscious self-interest, they are inevitably biased. Wilson (1989) notes,

> In defining a core mission [for an agency] and sorting out tasks that either fit or do not fit with this mission, executives must be aware of their many rivals for the right to define it. Operators with professional backgrounds will bring to the agency their skills, but also their biases: Lawyers, economists, and engineers see the world in very different ways. You cannot hire them as if they were tools that in your skilled hands will perform exactly the task you set for them. Black and Decker may make tools like that, but Harvard and MIT do not. Worker peer groups also set expectations to which operators conform, especially when the operators work in a threatening, unpredictable, or confrontational environment. (p. 371)

What if Harvard and MIT could make the type of tools Wilson is alluding to—that is, AI systems that are rational rather than self-interested, biased, influenced by their environment, or alternatively, programmed to apply only certain values and motives? Would machines with this capability improve the responsiveness of administrators by eliminating unwanted or unknown values or biases?

Indeed, Danziger, Dutton, Kling, and Kraemer (1982) note that Herbert Simon has seen computers as "an apolitical technology for improving the rationality of decisionmaking and the efficiency of operations in organizations" (p. xii). The AI community also engages in discussions of rationality and the nature of the mind. Franklin (1995) refers to one of the definitions of *mind* from the Oxford English Dictionary as "the cognitive or intellectual powers, as distinguished from the will and emotions. Often contrasted with heart" (p. 22). Franklin refers to this definition as describing what he thinks of as the "rational mind, a deliberative mind." By referring to this definition, Franklin raises the possibility of a machine that is wholly rational, that is, a technology that has the ability to separate facts from values, in the words of Herbert Simon.

However, the issue in the public sphere is not to remove values from decisions. We know that in public policy, there is no such thing as an apolitical or valueless decision; all but the most narrowly technical decisions reflect value choices or biases by what is done and by what is not done. What is intriguing from the AI field is the potential to know and predict underlying values and biases, thereby increasing the probability that decisions are made in a manner that is responsive to agreed-on premises (whose premises is a question we will return to later). The possibility of machines with programmed values is raised by Franklin (1995) in his description of the work of Ackley and Littman, who have developed artificial agents that both evolve and learn to deal with an environment that is both relatively dynamic and complex. This is done by building in an evaluation network, an artificial neural network, that "provides a mechanism by which values influence the learning of behavior" (Franklin, 1995, p. 206).

What is being examined within the AI community is therefore not traditional notions of machines without values but the ability to develop artificial systems that can be programmed with specific values, motives, and goals. Thus, there is the potential to have an AI system that approximates ideal rational man in the sense that all assumptions underlying the thinking and therefore decisionmaking processes can be understood.

From the administrative discretion perspective, such capabilities raise the possibility of an unquestioningly loyal system open to any set of values, motives, and goals that are imposed on it—the essence of the loyal career bureaucrat who serves the agenda of the current political administration without being influenced by personal interests, biases, or values.

The possibility of a machine with values brings us right back to a classic debate within public administration: What is the appropriate role for personal

values and interests in the ideal career public administrator? For those who favor the concept that democratic governance is served by programmed administrators (or intelligent machines) who respond dispassionately to elected officials and their appointees, then the possibilities provided by these systems is welcome. For those who are disturbed by the concept of the programmable administrator because of numerous examples in which administrators have exercised "loyalty that questions" and have curbed excesses of politicos, then the prospect of these systems is frightening.

Or perhaps there is some middle ground here where such a system could be useful for all parties as a tool for examining the underlying assumptions in any issue or decision. In other words, an AI system could be programmed with a variety of values, goals, and motives to see how different combinations would affect analyses or decisions.

JUDGMENT

Although there is disagreement over appropriate boundaries, most administrative discretion theorists acknowledge that public administrators must apply sound judgment in executing the law. In other words, no law can be written to cover all situations. A number of models have been developed to guide the exercise of good judgment. For example, Rohr (1986) presents the concept of autonomy grounded in subordination, which argues that public administrators must be open to the possibility that at times fidelity to their oath of office may require them to balance executive branch interests against other constitutional considerations. Dobel (1990) suggests that public administrators have multiple commitments to regime accountability, personal responsibility, and prudence. Individuals "of integrity should then iterate among the three realms in their judgements while using each other to balance and strengthen the others" (p. 354).

These and other models speak to the fact that there is a limit to how much one can program administration through the use of rigid rules. Ultimately, prudence or the ability to make judgments based on the situation is needed. Indeed, Wilson (1989) notes how street-level bureaucrats who interpret or otherwise bend the rules to fit the situation are most effective. He cites the following study of patrol officers by William K. Muir:

> The "good cops" were "street-corner politicians" who controlled their beats in the common interest by selectively enforcing the rules, sometimes letting off people for behavior for which others were arrested. The not-so-good cops were those who either retreated from the confusion and dangers of the street altogether or mechanically applied every rule as the law required. (Wilson, 1989, p. 344)

In other words, at some point bureaucrats inevitably must exercise prudence or judgment, and conventional wisdom conveys that this is where computers fall short and only humans can exercise such a capability. For example, Hummel (1994) notes that "without participation in the human experience, the computer is not capable of something like understanding" (p. 174). In their examination of payoffs from computerization over time, Northrup, Kraemer, Dunkle, and King (1990) raise a similar point:

> Some [computerization] payoffs may be realized only at a minimal level even after years of experience by the most technically advanced cities due to the political nature of the tasks. For example, payoffs from computerization for planning and management decisions are a minor part of these tasks given the often overriding influence of tuition, judgement and politics. (p. 506)

These statements reflect the traditional assumption that computers cannot think on their feet, so to speak, or learn to account for subjective factors that vary with a situation. After all, it is one thing to say that one can develop the ultimate rational machine by deciding which values, motives, or goals to input, but it is quite another to have a prudent machine—that is, an intelligence system that can be independently political.

However, increasingly sophisticated systems are being developed that will be sensitive to more subtle factors in the environment. For example, Wilson (1985) has set himself the task of producing an AI system called Animat that exhibits intelligent behavior. His concept of intelligence involves the "ability to be repeatedly successful in satisfying one's psychological needs in diverse, observably different, situations on the basis of past experience" (p. 16). Note that Wilson uses the term *psychological needs* rather than just *physical needs*, indicating a level beyond mere physical response and into the world of understanding changes in one's environment.

Related to the issue of judgment is what Crevier (1993) describes as the ability to change one's mind when circumstances require, that is, common sense. For example, a new generation of expert systems is equipped with a monitoring mechanism called a Truth Maintenance System, which is a knowledge base consisting of a set of possible exceptions to various statements (e.g., "Thou shalt not kill, except when . . . "). Crevier also cites the work of Janet Kolodner, whose modeling of the mind of former secretary of state Cyrus Vance for a doctoral project led to a technique called case-based reasoning. He explains,

> Kolodner's idea is to stop painstakingly trying to distill the knowledge of experts into rules and, instead, record it directly as the experts do, in the form of a series of well-documented cases. Confronted with a new case, such as a set of symptoms in a patient or the salient points of a legal argument, the computer would then search its knowledge banks for a similar case, adapt it to the new situation, and conclude accordingly. (Crevier, 1993, p. 239)

A final stream of AI research in developing artificial judgment is entitled Cyc (short for *encyclopedia*), a $25 million research project based on the conclusion by some researchers that

> no amount of finessing and fancy footwork would ever let a machine discover by itself such elementary facts as "Nothing can be in two places at once," or "Animals don't like pain," and "People live for a single solid interval of time". (Crevier, 1993, p. 240)

Humans need to know a colossal number of these common-sense assertions to get by in the world. This project is engaged in encoding millions of common-sense assertions or self-evident facts that humans learn starting at a very early age and that are never included in reference books.

Finally, the exercise of sound judgment requires not only common sense and experience but empathy or compassion as well. Do AI researchers expect to create machines with human feelings integrated with human thinking? Waldrop (1987) suggests the most honest answer to this question is "Who knows?" However, he points to a number of scientific developments that have brought the issue of emotion into sharper focus in the AI community. For example, he notes that

> armed with all the techniques of cellular and molecular biology, they have begun to clarify how our most basic emotions and drives arise at the molecular level. Hunger, thirst, pleasure, pain, depression, elation, anger, sexual arousal—all seem to be products of the chemistry of brain. (p. 130)

Waldrop also refers to AI research with the goal of enabling a computer to infer goals and plans of characters in a story—that is, subjective understanding. He suggests,

> In principle, there's no reason a computer couldn't use those same techniques to figure out when someone is sad or angry in real life. Such a computer could then be programmed to make the appropriate responses—saying comforting words in the first case, or moving the conversation toward a less provocative subject in the second. Indeed such a computer could even be said to "empathize" in some sense. (p. 132)

Waldrop's words have proved prophetic because AI systems are being developed today that can sense when its human counterpart is becoming frustrated and relate that to the problem the user wants to solve (Johnston, 1996; Mullich, 1999). Furthermore, a robot is being developed at MIT designed to have drives similar to human needs. The researcher's goal is that "the robot will learn that its expressions, and the intensity of those expressions, can cue its caregiver into specific actions" (Waldrop, 1987, p. 64). Such robots might be used in a variety of settings, such as in liaisons between patients and doctors.

Considering the usefulness of such technology in areas such as law enforcement, visions of the popular *Robocop* movies spring to mind, where ideal robotic police officers patrol the streets, scanning the environment and making independent, prudent decisions based on an acceptable range of values or motives (public safety, reasonable force) and not influenced by undesirable values (fear, hatred, self-interest, power).

Although such capabilities are mere science fiction today, the very fact that researchers are actively experimenting with this level of machine intelligence raises questions regarding what such machines would be taught. In other words, just as humans receive some sort of training before they are sent off on their jobs, what would we teach such machines embarking on roles in the field of public administration?

Although we have accredited professional degree programs in the field of public administration and public policy, we still struggle in the shadow of more mature fields such as medicine and law with identifying what Lynn (1997) refers to as the "professional reasoning processes" that distinguish professional work in the public sector. In other words, is there an identifiable way we want a public administrator in different settings to think? In a recent work on this topic, Lynn (1996) refers to Eliot Friedson's view that to claim professional status, a group must demonstrate the "work they do is esoteric, complex, and discretionary in character: it requires theoretical knowledge, skill, and judgement that ordinary people do not possess, may not wholly comprehend, and cannot readily evaluate" (p. 147).

Once again, although certainly not a new dilemma, the lens of AI reinforces the importance of the pursuit of the professionalism for the field of public administration. This pursuit is not important for the mere sake of proprietary control over certification processes but because it clarifies what it means to think like a professional public administrator, that is, to approach situations and decisions with a certain discipline and habits of mind.

ACCOUNTABILITY

Prior literature on computerization expressed concern over changes in balance in power, that is, that technical experts in control of computers would gain increased influence over elected officials and therefore increase the influence of unelected technocrats in the government system. To this point of computerization, such fears have not been justified; rather, computer use has tended to reinforce existing power relations in organizations. Danziger et al. (1982) note, that "computing tends to reinforce not only the prevailing structure of control within local governments, but also the prevailing political and organizational biases of those governments. In this sense, computing has been a politically conservative technological innovation" (p. 3).

Would the new level of computing represented by AI be potentially different? There are issues regarding control at two levels. First, would AI systems lessen the reliance of legislators on the bureaucracy, or vice versa (and if so, is this desirable?); and second, does the nature of AI (i.e., the ability to think independently) suggest systems that are potentially beyond the control of those responsible for them (i.e., public officials)?

In terms of the first question, it does not seem that the increased technical capacity potentially represented by an AI system would affect the balance of control between legislators or administrators other than the potential to reduce the number of analysts on both sides of the fence. As with the current situation with automation, all sides in the public policy arena (legislative staffs, executive branch analysts, interest groups), would match AI systems and have at it.

Furthermore, as Lowi (1969) argues, technology has not been shown to trump ideology in legislative matters. For example, what Lowi views as the inappropriate rise in administrative power is due to deliberately vague laws rather than "the usual cry of how complex and technological a new field is" (p. 144). He elaborates as follows:

> Delegation has been elevated to the highest of virtues, and standards have been relegated to the wastebasket of history because that is the logic of interest-group liberalism. Bargaining—or, as Schlesinger might call it, participation in the "interior processes of policy-making"—must be preferred over authority at every level and phase of government. (p. 144)

In other words, Lowi would not see technology as the issue. Vague legislation and the resulting delegation of power to career bureaucrats is not an accident or the result of inadequate technological capability and expertise in the legislative branch; it is a result of the ideology of interest group liberalism. Thus, once again new technology does not resolve an old dilemma.

However, perhaps what AI can potentially affect is the ability of citizens to hold public officials more accountable. In arguing for the importance of citizen involvement in achieving administrative responsibility, Cooper (1990) notes how Carl Friedrich "rightly predicted that the legislative task would become so enormous and complex that citizens would increasingly turn to the administrator to affect the workings of government" (p. 178). However, communication technology is having an impact on this equation, as seen by the increased ability of citizens to communicate with government officials through electronic mail and the potential for citizen education via the Internet. Nonetheless, even with the development of increasingly sophisticated search engines, the usefulness of the Internet relies on significant knowledge and time on the part of the user to find useful information. Citizens with access to AI systems with the ability to independently scan databases and learn about issues in a rapid manner may address this limitation and significantly improve the capability of citizens to

make independent assessments of public policy issues. AI technology is perhaps at least a partial answer to the dilemma over how citizens with limited time, expertise, and resources can assess decisions by public officials.

In terms of the second question regarding control, there is the distinct danger of AI systems becoming used so much that they become impossible to monitor either because they simply exceed the capacity of humans or the humans lose mastery of their subject area through lack of use. In other words, the danger here for the next generation that has access to machines that can independently think and make judgments is the temptation of relying too heavily on these machines. The issue here is more subtle than the obvious danger that machines may become malevolent or self-serving (as the computer Hal in the movie *2001: A Space Odyssey*). Clearly, the possibility of a malevolent computer is logically inescapable once one has a system that can learn independently, simply because one cannot control the type of information being absorbed by the computer, just as one cannot totally control a human's learning process.

However, even assuming benevolent AI systems, there is the danger of humans losing sight of the machine's operating assumptions, and we return once again to a classic problem in decision making: the failure to examine underlying assumptions. AI systems are an improvement on democratic governance (i.e., accountable decision making or exercise of discretion) only as long as the assumptions on which they are operating are known and under control. Otherwise, we are no better off, and perhaps worse off, than the current situation. We will simply have replaced unaccountable human administrators with unaccountable machines.

SUMMARIZING: POTENTIAL BENEFITS AND DANGERS OF AI FOR PUBLIC ADMINISTRATION

By using the lens of AI, this article has examined three classic concepts related to administrative discretion: responsiveness, judgment, and accountability. Future generations of public administrators will have access to technology that far surpasses our current concept of computers. Rather than high-speed calculators or even expert rule-based systems, scientists are developing AI systems that will not only learn to learn independently but could also possess values, motives, and goals. Furthermore, these systems may be capable of making subjective, political judgments.

The possibility of such systems in the hands of elected officials, public administrators, and the public at large forces us to examine a number of questions related to administrative responsiveness, reasoning processes unique to public administrators, and how we think about controlling or holding public servants accountable. This discussion points to a number of potential benefits as well as dangers for the field of public administration.

BENEFITS OF AI

Ability to Examine Underlying Assumptions and Values

One of the flaws in human decision making documented in a number of classic public policy cases is acting on bad assumptions, fear, or bias (Neustadt & May, 1986). What makes these conditions even worse is that these flaws may be unstated or even beyond the awareness of the individuals involved (Argyris & Schon, 1978). Even individuals with perfect information and awareness may ultimately lack the courage to speak up or take what they know is the proper action. Perhaps if presidents Kennedy or Roosevelt and their advisers had access to AI, mistakes such as the Bay of Pigs or the internment of Japanese American citizens may have been averted. Reflection by participants in both of these tragic cases reveals individuals who did not speak up or share information because of fear, intimidation, or hidden agendas (Barth, 1992; Neustadt & May, 1986). An AI system would be devoid of such baggage. Or on a less dramatic level, perhaps everyday citizens would be treated more fairly by representatives of a government equipped with AI technology. For example, Boden (1990) points out the advantages of an AI system that can be programmed for values:

> Many people—for instance, those who are female, working class, Jewish, disabled, or black—encounter unspoken (and often unconscious) prejudice in their dealings with official or professional bodies. An AI welfare advisor, for example, would not be prejudiced against such clients unless its data and inferential rules were biased in the relevant ways. A program could, of course, be written so as to embody its programmer's prejudices, but the program can be printed out and examined, whereas social attitudes cannot. (p. 451)

Thus, AI systems may enhance the opportunity for decision makers to have access to analyses incorporating explicit goals, motives, and values that at the same time are unclouded by human frailties such as anger, fear, or prejudice.

Identification of Reasoning Processes Unique to Professional Public Administrators

The specter of artificially intelligent "robotic" administrators provides an incentive to continue examining what it means to think like a public administrator. To illustrate, Tompkins, Laslovich, and Greene (1996) suggest that a competent public administrator is "one who is politically sophisticated, technically competent, well grounded in conceptual and theoretical knowledge, and cognizant of the values that connect the means and ends of government" (p. 121). Programming an AI system to be technically competent and well-grounded in concepts and theories is a task with relatively defined parameters. However, this challenge becomes particularly interesting when one considers how to program

a system to be politically sophisticated or to incorporate certain values. For example, what would a program say about balancing freedom versus order or efficiency versus responsiveness? How about the scenario of "always following the directives of your political superior, except when . . . " Or, following on the case-based reasoning project discussed earlier, what are the most useful cases in public administration (real or created) from which an AI system could learn best? Access to such an artificial mind could be a wonderful aid to a practicing public official, but simply the process of thinking through how to program such a system would also be a valuable learning experience for public administration education scholars.

In a similar vein, AI could improve the use of standardized aptitude tests such as the computer-adaptive Graduate Record Examination (GRE) for admission to public administration graduate schools. Although research in the business field is examining the use of AI techniques to predict admission decisions to business schools, the focus in these efforts is on more efficient and effective methods to evaluate existing criteria such as grade point average, Graduate Management Admissions Test (GMAT) scores, and professional references (Ragothaman & Davies, 1998). However, given the potential capabilities of AI systems described in this article, the usefulness of the actual standardized aptitude tests such as the GMAT or GRE could be improved upon. The chronic complaint about these aptitude tests is that they are culturally biased or too narrow in their focus because they only capture cognitive knowledge. Advances in AI technology may allow us to design tests more customized to particular fields or to real time evaluation of situations or cases. In theory, AI systems could allow for what is in effect a personal interview with test takers, walking individuals through case scenarios, placing individuals under various time and other stressors, or using other techniques that may allow test takers with different values, backgrounds, or abilities to come through. For example, it is possible that a nontraditional but experienced student who does not test well on traditional cognitive tests might perform quite well on a test that requires one to respond to a complex case scenario under pressure. Such a test would not necessarily replace the standardized, cognitive tests currently used but could be a useful supplement.

Enhancement of Citizen Knowledge

Perhaps the most potentially powerful benefit of AI is the availability of such technology to the citizenry. Assuming that, similar to most technology, artificially intelligent machines would be accessible to the average American over time, the potential for providing citizens with information and analyses previously unavailable could address the need for a more educated and sophisticated citizenry. The ability for the average American, with constrained time and energy, to truly understand the ramifications of various public policy debates is limited at best. An AI system, with the ability to learn and draw on the vast databases on the Internet as well as other government databases open to public

access, could walk a citizen through an issue, answering questions, responding to different scenarios, and so forth. In the business field, this concept is known as data mining, a multibillion dollar industry based on developing software that can efficiently extract useful information from massive databases and use it as a basis for making business decisions (Cory, 1999). With similar capabilities to mine data relevant to public policy issues, citizens in the future may be able to consult their AI political adviser rather than rely on sound bites, negative campaign ads, and debates in the newspaper and on television where complex issues are oversimplified.

For the public administrator, particularly at the local level, citizens and interest groups armed with AI advisers could very well level the playing field with regard to expertise; the administrator may no longer be able to claim so readily that he or she knows best, particularly if the citizens' AI systems have access to the same databases. For example, faced with a complex annexation or tax issue, a citizen could use an AI system to walk though the relevant laws and ordinances, assess the impacts of different scenarios given an array of goals or values, and essentially learn along with the AI system. Instead of relying totally on the information provided by lawyers and bureaucrats, citizens could have an enhanced understanding of issues and, therefore, the ability to ask better questions.

DANGERS OF AI

Who Is Doing the Programming?

A vexing problem is created by computers that learn independently and exercise judgement. With static rule-based systems, it is relatively simple to identify the program being used and examine the assumptions built in. However, AI systems that can learn independently place the original programmer in an even more vital position. Whitby (1984) suggests that as the development of AI programs becomes less an academic pursuit and more a matter of direct commercial and government interest, the need for security and specialization will grow and create a situation where elites dominate. He warns

> At present the main force mitigating secrecy in commercial and military "stupid" computer systems is the perceived danger of being too much at the mercy of any single human software designer. The human threat alone prompts documentation of present "stupid" military systems. Programs capable of self-development afford a most convenient way of avoiding this human threat to security. The commercial and military paymasters will be constrained by economic and competitive forces to concentrate on questions of whether or not a system works, rather than spending time and money finding out how it works. . . . They will not, unless something is done about it now, be detained by the need for perspicuous documentation. . . . In particular, the practitioners of commercial and military AI will be increasingly

> treated as a highly-rewarded specialist group who will be discouraged from revealing the secrets of their craft to outsiders. (pp. 239-240)

The realm of public administration is an excellent example of the potential danger of elitism because of the fundamental importance of accountability. Citizens expect government officials to make decisions in the public interest based on the best information available. To ensure such behavior, officials can be taken to task in public hearings, congressional inquiries, and the examination of public records. Furthermore, we expect decisions to be made by individuals whose qualifications and expertise are apparent; that is, they have a vita open to public inspection. Put another way, we expect and can confirm that our city's chief urban planner is not a gerontologist by training, that he or she has the appropriate graduate degree, and so forth. What is the equivalent of a vita for an AI system? Logically, the equivalent would have to be their programmers, simply because the only way to assess the quality of the information produced by an AI system is the premises on which the machine makes its decisions. As Herbert Simon explained, the key to understanding decision making is to understand the premises; it can be no different with AI systems. There will have to be very tight controls or regulations placed on the development of AI systems with the capacity to learn, very much like we currently have with accreditation and credential requirements for universities. If we are going to have AI systems that approach independent human thinking but in fact are even more powerful because of the vast databases, speed, and limitless attention spans possessed by these systems, we must think about how we assess the quality of individual artificial minds, just as we attempt to assess the quality of human talent in our public organizations.

A Less Representative Bureaucracy

A final point related to the elitist programming problem as it applies to a government setting is the importance of a diverse and representative government bureaucracy. During a particularly activist time for the Federal Reserve Board during the Carter administration, one heard the cry "Who elected Paul Volcker?" referring to concern over the powerful yet unaccountable influence of the appointed Federal Reserve Board chairman. This sentiment reflects concern about decisions made by administrators beyond the reach of voters. A traditional response to this concern over the legitimacy and responsiveness of the nonelected bureaucracy or administrative state is that its members are representative of the general population; that is, we have public employees of all backgrounds involved in the analysis and implementation of public policy. For example, Rohr (1986) suggests that the modern administrative state, with its involvement of millions of ordinary citizens in the systematic execution of public law, provides a balance or check against "the likely excesses of a single executive prone to carry out his constitutional powers in a haughty or arrogant manner that offends republican principle" (p. 48).

Following the path of automation in history, AI will likely reduce the need for the number of analyst positions in the government. Combined with the potentially elitist nature of the parents (i.e., programmers) of the AI systems replacing these human (and presumably more diverse) public servants, the threat to the representative bureaucracy argument is apparent. Logically, if we accept the impact of the parents on the views of children, we must accept the imprinting that occurs on artificial minds by their potentially elitist programmers as well.

Atrophying of Administrators' Own Judgment and Sense of Responsibility

Another major stream of caution about AI is the direct relationship between the increasing capabilities of artificially intelligent systems and the tendency for humans to rely on them. This has always been a concern with computerization and is reflected in scenarios such as, "I'm sorry Mr. Governor (or Mr. Mayor . . .), the computer doesn't allow us to do that" (Yazdani, 1986). Or perhaps even more chillingly, is the specter of public administrators unwilling to consider concerns raised by citizens because the computer does not share this concern? As long as the computer is only viewed as a tool or an aid, this fear has been contained. However, the new level of computing represented by AI may change this equation. Machines that cannot only go beyond the application of direct rules and learn independently but can also communicate with users in their native language (another aspect of emerging AI technology) may be very difficult for humans to question. Yazdani notes the tendency of human beings "to attribute programs far more intelligence than they actually possess (by any reasonably objective measure) as soon as the program communicates in English [or other natural language] phrases" (p. 326).

Such concerns are not trivial, as demonstrated by the resignation of David Parnas, a highly respected computer scientist, from the U.S. government's top advisory committee on the Strategic Defense Initiative because of what he believed to be an inappropriate reliance on AI technology (Boden, 1990). Furthermore, Shneiderman (1999) argues that enhancing computers' autonomy raises troubling questions about who will be responsible if AI systems controlling air traffic or medical equipment, for example, make errors that end in disaster.

To foster accountability, responsible public administrators need to practice what Hugh Heclo calls "loyalty that argues back" (Cooper, 1990). In other words, it is one thing to independently think through and analyze a question posed by one's superior; however, it is quite another to raise additional unasked questions or take the initiative and provide unsolicited advice where warranted. Even the most sophisticated AI system ultimately may be flawed because it lacks curiosity—that is, the urge to investigate issues or ask questions on its own. Shneiderman (1999) is concerned that coupling humans with AI systems may hinder human potential, because people are potentially "richly creative and generative in ways that amaze me and that defy simple modeling" (p. 35). We

expect such behavior in responsible government officials. AI systems can be a great benefit as advisers or tools to human officials but a danger if their dazzling capabilities create overly passive or deferential human counterparts.

CONCLUSION

Returning to the original question raised by the MPA student at the beginning of this article—that is, whether the emergence of AI makes administrative discretion arguments moot—the answer must be a resounding no. In fact, advances in AI technology make these discussions more important than ever. The real danger of AI in government is represented by researchers who are divorced from the world of public administration scholars and practitioners and are engaged in discussions and making technological decisions without understanding the implications for governance of the administrative state. Future generations would be well served by such reflection because history demonstrates the turmoil that results from considering the full social and political implications of new technology only after it is upon us. Boden (1990) provides an appropriate closing sentiment on the implications of this technology:

> Is artificial intelligence in human society a utopian dream or a Faustian nightmare. . . . If future generations are to have reason to thank us rather than to curse us, it's important that the public (and politicians) of today should know as much as possible about the potential effects—for good or ill—of artificial intelligence. (p. 450)

REFERENCES

Ansley, L. (1994, February 6). The doctor is online. *USA Weekend*, p. 26.

Argyris, C., & Sclon, D.A. (1978). *Organizational learning: A theory of action perspective.* Reading, MA: Addison-Wesley.

Barth, T. J. (1992). The public interest and administrative discretion. *American Review of Public Administration, 22*(4), 289-300.

Boden, M. A. (1990). The social impact of artificial intelligence. In R. Kurzweil (Ed.), *The age of intelligent machines* (pp. 450-453). Cambridge: Massachusetts Institute of Technology.

Cook, B. (1995). Breaking chains or forging them? The debate on the scope of administrative discretion. Unpublished manuscript.

Cooper, T. (1990). *The responsible administrator.* San Francisco: Jossey-Bass.

Cory, J. (1999). Software does the thinking. *Philadelphia Business Journal, 17*(51), B1-B2.

Crevier, D. (1993). *AI: The tumultuous history of the search for artificial intelligence*. New York: Basic Books.

Daily Defense News Capsules. (1995, June 27). Information Access Company, a Thompson Corporation Company IAC (SM) Newsletter Database.

Danziger, J. N., Dutton, W. H., Kling, R., & Kraemer, K. (1982). *Computers and politics*. New York: Columbia University.

Dobel, P. J. (1990, May/June). Integrity in the public service. *Public Administration Review,* pp. 354-366.

Dreyfus, H. L. (1994). *What computers still can't do*. Cambridge, MA: MIT Press.
Franklin, S. P. (1995). *Artificial minds*. Cambridge, MA: MIT Press.
Gill, K. S. (1986). *Artificial intelligence for society.* New York: John Wiley.
Grupe, F. H., von Sadovsky, T., Owrang O., Mehdi, M. (1995). An executive's guide to artificial intelligence. *Information Strategy: The Executive's Journal*, *12*(1), 44.
Hummel, R. P. (1994). *The bureaucratic experience*. New York: St. Martin's Press.
Hunter, C. (1995, November 16). Maybe it's not real "top gun" but it's close. *Hartford Courant*, p. E3.
Johnston, S. J. (1996). IBM improving support for warp. *Information Week*, *583*, 99.
Kurzweil, R. (1990). *The age of intelligent machines*. Cambridge: Massachusetts Institute of Technology.
Kurzweil, R. (1999). *The age of spiritual machines*. New York: Viking.
Lowi, T. J. (1969). *The end of liberalism*. New York: Norton.
Lowi, T. J. (1993). Legitimizing public administration: A disturbed dissent. *Public Administration Review*, *53*(3), 261-264.
Lynn, L. E., Jr. (1996). *Public management as art, science, and profession*. Chatham, NJ: Chatham House.
Lynn, L. E., Jr. (1997, September 25-27). Remarks at the 1997 Southeastern Conference on Public Administration, Knoxville, TN.
Miller, D. (1994). French carrier alone in its class. *International Defense Review*, *27*, 59.
Mullich, J. (1999). Give your computer a hanky. *Business Journal*, *13*(4), 27.
Neustadt, R.E., & May, E.R. (1986). *Thinking in time: The uses of history for decision-makers*. New York: Free Press.
Northrup, A., Kraemer, K. L., Dunkle, D., & King, J. L. (1990). Payoffs from computerization: Lessons over time. *Public Administration Review*, *50*(5), 505-514.
Pappas, B. (1998). Transparent eyeball. *Forbes*, *162*(14), 47.
Pech, Janet (1999). Scudder launches artificial intelligence on new site. *Financial NetNews*, *4*(6), 1.
Peterson, I. (1993). AI at work: From prisons to autos to space. *Science News*, *144*(3), 37.
Raghupathi, W. (1991). Exploring connectionist approaches to legal decisionmaking. *Behavioral Science*, *36*(2), 133-139.
Ragothaman, S., Davies, T., et al. (1998). Using artificial intelligence techniques to predict M.B.A. admission decisions. *College Student Journal*, *32*(1), 125-135.
Rohr, J. A. (1986). *To run a constitution*. Lawrence: University of Kansas.
Sack, W. (1997). Artificial human nature. *Design Issues*, *13*(2), 55-65.
Shneiderman, B. (1999). Humans unite! *Scientific American*, *280*(3), 35-37.
Seibel, T. (1994, July 4). Computer profiles to help sniff out crooked cops. *Chicago Sun-Times*, p. 4.
Southerst, J. (1992). An arm with a mind of its own. *Canadian Business*, *65*(3), 60-64.
Spagnoletti, R. J. (1987, Winter). Using artificial intelligence to aid in the resolution of socioscientific disputes: A theory for the next generation. *Journal of Law and Technology.*
Suplee, C. (1997). Robot revolution. *National Geographic*, *192*, 76-95.
Tompkins, J., Laslovich, M. J., & Greene, J. D. (1996). Developing a competency-based MPA curriculum. *Journal of Public Administration Education*, *2*(2), 117-130.
Trefil, J. (1997). *Are we unique: A scientist explores the unparalleled intelligence of the human mind*. New York: John Wiley.
Waldrop, M. M. (1987). *Man-made minds*. Lisle, IL: Walker.
Whitby, B. (1984). AI: some immediate dangers. In M. Yazdani & A. Narayanan (Eds.), *Artificial intelligence: Human effects* (pp. 235-245). New York: Halstead.
Wilson, J. Q. (1989). *Bureaucracy.* New York: Basic Books.
Wilson, S. W. (1985). Knowledge growth in an artificial animal. In *Proceedings of the first international conference on genetic algorithms and their applications* (pp. 16-23). Hillsdale, NJ: Lawrence Erlbaum.
Winston, P. H. (1992). *Artificial intelligence.* Reading, MA: Addison-Wesley.

World Wide Web http://www.yahoo.com.business-and-economy/companies/computers/software/ artificialintelligence, June 1998
Yazdani, M. (1986). *Artificial intelligence: Principles and applications*. New York: Chapman and Hall.
Yazdani, M., & Narayanan, A. (1984). *Artificial intelligence: Human effects*. Chichester, UK: Ellis Horwood.

Thomas J. Barth is an associate professor in the Department of Political Science at The University of North Carolina at Wilmington. He is currently planning a new MPA program.

Eddy Arnold is the chief communication engineer in the Department of Communications at the University of Memphis. He recently earned an MPA degree.