AI and Its Impact on Public Administration

Standing Panel on Technology Leadership
April 2019

NATIONAL ACADEMY OF PUBLIC ADMINISTRATION®

The National Academy of Public Administration (the Academy) is an independent, nonprofit, and nonpartisan organization established in 1967 to assist government leaders in building more effective, accountable, and transparent organizations. Chartered by Congress to provide nonpartisan expert advice, the Academy's unique feature is its over 900 Fellows—including former cabinet officers, Members of Congress, governors, mayors, and state legislators, as well as prominent scholars, business executives, and public administrators. The Academy helps the federal government address its critical management challenges through in-depth studies and analyses, advisory services and technical assistance, congressional testimony, forums and conferences, and online stakeholder engagement. Under contracts with government agencies, some of which are directed by Congress, as well as grants from private foundations, the Academy provides insights on key public management issues, as well as advisory services to government agencies.

Copyright © 2019 by National Academy of Public Administration.
All rights reserved. Published and hosted by the Academy.

ARTIFICIAL INTELLIGENCE AND ITS IMPACT ON PUBLIC ADMINISTRATION

NATIONAL ACADEMY OF PUBLIC ADMINISTRATION
STANDING PANEL ON TECHNOLOGY LEADERSHIP'S WORKING GROUP ON ARTIFICIAL INTELLIGENCE AND ROBOTICS AND THE IMPACT ON PUBLIC ADMINISTRATION

Organized and edited by Alan R. Shark

Authors
(Listed in order of appearance):
Karen Shrum
Lisa Gordon
Priscilla Regan
Karl Maschino
Alan R. Shark
Anders Shropshire

Table of Contents

Contributing Fellows and Experts

This series of papers would not have been possible without the collected wisdom and contributions of many people. We are especially indebted to this group for their challenging questions, brainstorming, dedicated effort, and detailed review that made this work substantially better.

Robert Bland
Marjory Blumenthal
Elliot Branch
Dan Chenok
Elizabeth Fretwell
Lisa Gordon
Bill Greenwalt
Jim Hendler
David Horn
Karl Maschino
Priscilla Regan
Dave Rejeski
Alan Shark
Myra Shiplett
Karen Shrum
Mariko Silver
Teresa Takai
Carl Van Horn
Michael Van Milligen

Introduction

In early 2018, the Board of the National Academy of Public Administration (the Academy) challenged its Standing Panels to develop public facing activities that advance the mission of the Academy. Accepting this challenge, the Standing Panel on Technology Leadership formed the Artificial Intelligence (AI) Working Group, based on the strong belief that AI and Robotics are generating a huge amount of interest and have the potential to impact the field of public administration for many years to come. Further, the group believes that the Academy has a great opportunity to actively explore this subject area and seek meaningful outcomes.

In March 2018, the AI Working Group focused on how to best match the talents of Academy Fellows and associated volunteers with the goal of developing White Papers in areas that the group felt had the most potential to impact public administration. By the Spring of 2018, three working groups emerged that focused on:

- AI and The Future of Work
- AI and Ethics
- AI and the Public Administration Curriculum

The Working Group on the Future of Work was headed by Karen Shrum, Principal, Government and Public Sector, Ernst & Young, LLP (EY). The Working Group on AI and Ethics was headed by Priscilla M. Regan, Professor of Government and Politics, George Mason University's Schar School of Policy and Government. Finally, the Working Group on AI and the Teaching of Public Administration was headed by Alan R. Shark, Chair, Standing Panel on Technology Leadership; Executive Director, Public Technology Institute; and Associate Professor, the Schar School, George Mason University.

Each of the working groups are indebted to the work of Anders Shropshire, a Public Administration student at the University of Wisconsin recruited by the Academy to serve as a summer intern, as the group's principal researcher and, in many cases, its principal writer. His research, quality of writing, and grasp of the topics was nothing short of extraordinary.

It is the hope of all those involved in the visioning, thinking, research, and commenting on draft reports that the final white papers are accepted as topics for action, that each addresses practical concerns, and that each calls for on-going research and issue monitoring. Making this report even more timely is the fact that President Trump issued an *Executive Order on Maintaining American Leadership in Artificial* on February 11th, 2019. We have included a copy of this significant order in this publication so that the reader has easy reference to the Administration's AI objectives and a framework for our proposals.

The Working Groups believe their work has just begun, and we can expect more Academy contributions and outcomes on this overall topic in the years ahead.

Dr. Alan R. Shark
Chair, Standing Panel on Technology Leadership

Executive Order on Maintaining American Leadership in Artificial Intelligence

Issued February 11, 2019

By the authority vested in me as President by the Constitution and the laws of the United States of America, it is hereby ordered as follows:

Section 1. Policy and Principles. Artificial Intelligence (AI) promises to drive growth of the United States economy, enhance our economic and national security, and improve our quality of life. The United States is the world leader in AI research and development (R&D) and deployment. Continued American leadership in AI is of paramount importance to maintaining the economic and national security of the United States and to shaping the global evolution of AI in a manner consistent with our Nation's values, policies, and priorities. The Federal Government plays an important role in facilitating AI R&D, promoting the trust of the American people in the development and deployment of AI-related technologies, training a workforce capable of using AI in their occupations, and protecting the American AI technology base from attempted acquisition by strategic competitors and adversarial nations. Maintaining American leadership in AI requires a concerted effort to promote advancements in technology and innovation, while protecting American technology, economic and national security, civil liberties, privacy, and American values and enhancing international and industry collaboration with foreign partners and allies. It is the policy of the United States Government to sustain and enhance the scientific, technological, and economic leadership position of the United States in AI R&D and deployment through a coordinated Federal Government strategy, the American AI Initiative (Initiative), guided by five principles:

(a) The United States must drive technological breakthroughs in AI across the Federal Government, industry, and academia in order to promote scientific discovery, economic competitiveness, and national security.

(b) The United States must drive development of appropriate technical standards and reduce barriers to the safe testing and deployment of AI technologies in order to enable the creation of new AI-related industries and the adoption of AI by today's industries.

(c) The United States must train current and future generations of American workers with the skills to develop and apply AI technologies to prepare them for today's economy and jobs of the future.

(d) The United States must foster public trust and confidence in AI technologies and protect civil liberties, privacy, and American values in their application in order to fully realize the potential of AI technologies for the American people.

(e) The United States must promote an international environment that supports American AI research and innovation and opens markets for American AI industries, while

protecting our technological advantage in AI and protecting our critical AI technologies from acquisition by strategic competitors and adversarial nations.

Sec. 2. Objectives. Artificial Intelligence will affect the missions of nearly all executive departments and agencies (agencies). Agencies determined to be implementing agencies pursuant to section 3 of this order shall pursue six strategic objectives in furtherance of both promoting and protecting American advancements in AI:

(a) Promote sustained investment in AI R&D in collaboration with industry, academia, international partners and allies, and other non-Federal entities to generate technological breakthroughs in AI and related technologies and to rapidly transition those breakthroughs into capabilities that contribute to our economic and national security.

(b) Enhance access to high-quality and fully traceable Federal data, models, and computing resources to increase the value of such resources for AI R&D, while maintaining safety, security, privacy, and confidentiality protections consistent with applicable laws and policies.

(c) Reduce barriers to the use of AI technologies to promote their innovative application while protecting American technology, economic and national security, civil liberties, privacy, and values.

(d) Ensure that technical standards minimize vulnerability to attacks from malicious actors and reflect Federal priorities for innovation, public trust, and public confidence in systems that use AI technologies; and develop international standards to promote and protect those priorities.

(e) Train the next generation of American AI researchers and users through apprenticeships; skills programs; and education in science, technology, engineering, and mathematics (STEM), with an emphasis on computer science, to ensure that American workers, including Federal workers, are capable of taking full advantage of the opportunities of AI.

(f) Develop and implement an action plan, in accordance with the National Security Presidential Memorandum of February 11, 2019 (Protecting the United States Advantage in Artificial Intelligence and Related Critical Technologies) (the NSPM) to protect the advantage of the United States in AI and technology critical to United States economic and national security interests against strategic competitors and foreign adversaries.

Sec. 3. Roles and Responsibilities. The Initiative shall be coordinated through the National Science and Technology Council (NSTC) Select Committee on Artificial Intelligence (Select Committee). Actions shall be implemented by agencies that conduct foundational AI R&D, develop and deploy applications of AI technologies, provide educational grants, and regulate

and provide guidance for applications of AI technologies, as determined by the co-chairs of the NSTC Select Committee (implementing agencies).

Sec. 4. Federal Investment in AI Research and Development.

(a) Heads of implementing agencies that also perform or fund R&D (AI R&D agencies), shall consider AI as an agency R&D priority, as appropriate to their respective agencies' missions, consistent with applicable law and in accordance with the Office of Management and Budget (OMB) and the Office of Science and Technology Policy (OSTP) R&D priorities memoranda. Heads of such agencies shall take this priority into account when developing budget proposals and planning for the use of funds in Fiscal Year 2020 and in future years. Heads of these agencies shall also consider appropriate administrative actions to increase focus on AI for 2019.

(b) Heads of AI R&D agencies shall budget an amount for AI R&D that is appropriate for this prioritization.

(i) Following the submission of the President's Budget request to the Congress, heads of such agencies shall communicate plans for achieving this prioritization to the OMB Director and the OSTP Director each fiscal year through the Networking and Information Technology Research and Development (NITRD) Program.

(ii) Within 90 days of the enactment of appropriations for their respective agencies, heads of such agencies shall identify each year, consistent with applicable law, the programs to which the AI R&D priority will apply and estimate the total amount of such funds that will be spent on each such program. This information shall be communicated to the OMB Director and OSTP Director each fiscal year through the NITRD Program.

(c) To the extent appropriate and consistent with applicable law, heads of AI R&D agencies shall explore opportunities for collaboration with non-Federal entities, including: the private sector; academia; non-profit organizations; State, local, tribal, and territorial governments; and foreign partners and allies, so all collaborators can benefit from each other's investment and expertise in AI R&D.

Sec. 5. Data and Computing Resources for AI Research and Development.

(a) Heads of all agencies shall review their Federal data and models to identify opportunities to increase access and use by the greater non-Federal AI research community in a manner that benefits that community, while protecting safety, security, privacy, and confidentiality. Specifically, agencies shall improve data and model inventory documentation to enable discovery and usability, and shall prioritize improvements to access and quality of AI data and models based on the AI research community's user feedback.

(i) Within 90 days of the date of this order, the OMB Director shall publish a notice in the *Federal Register* inviting the public to identify additional requests for access or

quality improvements for Federal data and models that would improve AI R&D and testing. Additionally, within 90 days of the date of this order, OMB, in conjunction with the Select Committee, shall investigate barriers to access or quality limitations of Federal data and models that impede AI R&D and testing. Collectively, these actions by OMB will help to identify datasets that will facilitate non-Federal AI R&D and testing.

(ii) Within 120 days of the date of this order, OMB, including through its interagency councils and the Select Committee, shall update implementation guidance for Enterprise Data Inventories and Source Code Inventories to support discovery and usability in AI R&D.

(iii) Within 180 days of the date of this order, and in accordance with the implementation of the Cross-Agency Priority Goal: Leveraging Federal Data as a Strategic Asset, from the March 2018 President's Management Agenda, agencies shall consider methods of improving the quality, usability, and appropriate access to priority data identified by the AI research community. Agencies shall also identify any associated resource implications.

(iv) In identifying data and models for consideration for increased public access, agencies, in coordination with the Senior Agency Officials for Privacy established pursuant to Executive Order 13719 of February 9, 2016 (Establishment of the Federal Privacy Council), the heads of Federal statistical entities, Federal program managers, and other relevant personnel shall identify any barriers to, or requirements associated with, increased access to and use of such data and models, including:

(A) privacy and civil liberty protections for individuals who may be affected by increased access and use, as well as confidentiality protections for individuals and other data providers;

(B) safety and security concerns, including those related to the association or compilation of data and models;

(C) data documentation and formatting, including the need for interoperable and machine-readable data formats;

(D) changes necessary to ensure appropriate data and system governance; and

(E) any other relevant considerations.

(v) In accordance with the President's Management Agenda and the Cross-Agency Priority Goal: Leveraging Data as a Strategic Asset, agencies shall identify opportunities to use new technologies and best practices to increase access to and usability of open data and models, and explore appropriate controls on access to sensitive or restricted data and models, consistent with applicable laws and policies, privacy and confidentiality protections, and civil liberty protections.

(b) The Secretaries of Defense, Commerce, Health and Human Services, and Energy, the Administrator of the National Aeronautics and Space Administration, and the Director of the National Science Foundation shall, to the extent appropriate and consistent with applicable law, prioritize the allocation of high-performance computing resources for AI-related applications through:

(i) increased assignment of discretionary allocation of resources and resource reserves; or

(ii) any other appropriate mechanisms.

(c) Within 180 days of the date of this order, the Select Committee, in coordination with the General Services Administration (GSA), shall submit a report to the President making recommendations on better enabling the use of cloud computing resources for federally funded AI R&D.

(d) The Select Committee shall provide technical expertise to the American Technology Council on matters regarding AI and the modernization of Federal technology, data, and the delivery of digital services, as appropriate.

Sec. 6. Guidance for Regulation of AI Applications.

(a) Within 180 days of the date of this order, the OMB Director, in coordination with the OSTP Director, the Director of the Domestic Policy Council, and the Director of the National Economic Council, and in consultation with any other relevant agencies and key stakeholders as the OMB Director shall determine, shall issue a memorandum to the heads of all agencies that shall:

(i) inform the development of regulatory and non-regulatory approaches by such agencies regarding technologies and industrial sectors that are either empowered or enabled by AI, and that advance American innovation while upholding civil liberties, privacy, and American values; and

(ii) consider ways to reduce barriers to the use of AI technologies in order to promote their innovative application while protecting civil liberties, privacy, American values, and United States economic and national security.

(b) To help ensure public trust in the development and implementation of AI applications, OMB shall issue a draft version of the memorandum for public comment before it is finalized.

(c) Within 180 days of the date of the memorandum described in subsection (a) of this section, the heads of implementing agencies that also have regulatory authorities shall review their authorities relevant to applications of AI and shall submit to OMB plans to achieve consistency with the memorandum.

(d) Within 180 days of the date of this order, the Secretary of Commerce, through the Director of the National Institute of Standards and Technology (NIST), shall issue a plan

for Federal engagement in the development of technical standards and related tools in support of reliable, robust, and trustworthy systems that use AI technologies. NIST shall lead the development of this plan with participation from relevant agencies as the Secretary of Commerce shall determine.

(i) Consistent with OMB Circular A-119, this plan shall include:

(A) Federal priority needs for standardization of AI systems development and deployment;

(B) identification of standards development entities in which Federal agencies should seek membership with the goal of establishing or supporting United States technical leadership roles; and

(C) opportunities for and challenges to United States leadership in standardization related to AI technologies.

(ii) This plan shall be developed in consultation with the Select Committee, as needed, and in consultation with the private sector, academia, non-governmental entities, and other stakeholders, as appropriate.

Sec. 7. AI and the American Workforce.

(a) Heads of implementing agencies that also provide educational grants shall, to the extent consistent with applicable law, consider AI as a priority area within existing Federal fellowship and service programs.

(i) Eligible programs for prioritization shall give preference to American citizens, to the extent permitted by law, and shall include:

(A) high school, undergraduate, and graduate fellowship; alternative education; and training programs;

(B) programs to recognize and fund early-career university faculty who conduct AI R&D, including through Presidential awards and recognitions;

(C) scholarship for service programs;

(D) direct commissioning programs of the United States Armed Forces; and

(E) programs that support the development of instructional programs and curricula that encourage the integration of AI technologies into courses in order to facilitate personalized and adaptive learning experiences for formal and informal education and training.

(ii) Agencies shall annually communicate plans for achieving this prioritization to the co-chairs of the Select Committee.

(b) Within 90 days of the date of this order, the Select Committee shall provide recommendations to the NSTC Committee on STEM Education regarding AI-related educational and workforce development considerations that focus on American citizens.

(c) The Select Committee shall provide technical expertise to the National Council for the American Worker on matters regarding AI and the American workforce, as appropriate.

Sec. 8. Action Plan for Protection of the United States Advantage in AI Technologies.

(a) As directed by the NSPM, the Assistant to the President for National Security Affairs, in coordination with the OSTP Director and the recipients of the NSPM, shall organize the development of an action plan to protect the United States advantage in AI and AI technology critical to United States economic and national security interests against strategic competitors and adversarial nations.

(b) The action plan shall be provided to the President within 120 days of the date of this order, and may be classified in full or in part, as appropriate.

(c) Upon approval by the President, the action plan shall be implemented by all agencies who are recipients of the NSPM, for all AI-related activities, including those conducted pursuant to this order.

Sec. 9. Definitions. As used in this order:

(d) the term "artificial intelligence" means the full extent of Federal investments in AI, to include: R&D of core AI techniques and technologies; AI prototype systems; application and adaptation of AI techniques; architectural and systems support for AI; and cyberinfrastructure, data sets, and standards for AI; and

(e) the term "open data" shall, in accordance with OMB Circular A-130 and memorandum M-13-13, mean "publicly available data structured in a way that enables the data to be fully discoverable and usable by end users."

Sec. 10. General Provisions.

(a) Nothing in this order shall be construed to impair or otherwise affect:

(i) the authority granted by law to an executive department or agency, or the head thereof; or

(ii) the functions of the Director of OMB relating to budgetary, administrative, or legislative proposals.

(b) This order shall be implemented consistent with applicable law and subject to the availability of appropriations.

(c) This order is not intended to, and does not, create any right or benefit, substantive or procedural, enforceable at law or in equity by any party against the United States, its departments, agencies, or entities, its officers, employees, or agents, or any other person.

DONALD J. TRUMP
THE WHITE HOUSE,
February 11, 2019.

Artificial Intelligence and the Future of Work

Prepared by Karen Shrum, Subgroup on the Future of Work Chair, and Lisa Gordon

Introduction

Today's working world is changing faster than ever before. Disruption is the new norm. For the first time we have four generations in the workforce: Traditionalist, Boomer, Gen X and Millennial. In addition, the definition of the employee has changed with essentially four types of workers: full-time, contingent, remote, and robot. Customers and employees are seeking the same digital experience in their work that they have at home. Automation is becoming an accepted part of work processes, with promises of efficiency and an enhanced focus on mission-oriented activities. As the speed of innovation accelerates, Artificial Intelligence (AI) will continue to be a factor driving change in how people work. It has the potential to reduce cost and increase the value delivered by the worker to the organization as it improves quality and speed of getting work done. It can deliver an anytime/anywhere digital and customer experience at work. It can free time and resources to allow workers to focus on strategic, high-touch activities.

The potential for increased productivity and work, as well as a blended workforce, can provoke both excitement and anxiety. For many, considering the benefits of productivity and increased quality, the choice to automate seems obvious. However, to others, the introduction of AI often brings trepidation and resistance, primarily around the loss of a job or a significant change in work. Leaders and managers may not know how to begin to implement AI as there are no readily discernable toolkits that provide a blueprint for successful implementation and adoption. These factors can impede and potentially stall the adoption of AI.

The adoption of AI in various business sectors is rapidly emerging, but is still in its early stages in government and not-for-profits. However, these sectors are not immune to the change and disruption being driven through AI. For example, the non-profit industry may be pressured to use AI by funders where they expect philanthropic dollars to be used wisely. The current trends usually focus on the outcomes of programs for clients, but in the future, there may be systems to automate how the work is done. Partnering across non-profit organizations, to track data for individuals who use services from multiple providers, is becoming a leading practice in many areas of the public sector. This collaborative approach and AI will likely converge in the future and lead to new service delivery models. The optimism to use AI is balanced by the need to continue to have a compassionate provision of service both to those in need and to the workers displaced by the introduction of new approaches to delivering services.

The intent of this paper is to help create awareness about AI's impact on the workforce and some of the resources that can inform leaders and managers as they prepare to deploy AI in their organizations. We look specifically at the federal, state and local and not-for-profit workforces. To set the stage, we begin with a brief overview of selected research. We summarize the key workforce challenges and what various organizations have done to overcome those as they implement AI. Based on our observations, we also present a road map comprised of practical

and actionable suggestions. It outlines how an organization may successfully prepare its workforce for AI deployment and ultimately provide better service delivery to its customers.

What the research tells us

Much has been written about AI and its impact on the workforce, but there is no consensus on how many jobs will be replaced entirely and what types of new jobs will be created. We selected a few examples for this paper: A recent World Economic Forum (WEF) report[1], *The Future of Jobs 2018*, noted that developments in automation technologies and AI could see 75 million jobs displaced. However, another 133 million new roles may emerge as companies shake up their division of labor between humans and machines, translating to 58 million net new jobs being created by 2022. Companies, governments and employees need to work together to tackle skills shortages and dislocations that occur due to automation, according to WEF.

Similarly, Ernst & Young LLP (EY) undertook proprietary research aimed at providing actionable insights to leaders ready to take advantage of automation and effectively drive business transformation. As part of this effort, they mapped Frey & Osborne automation scores to nearly 2,000 occupations in four economies (US, UK, Canada and Australia).[2] The EY researchers categorized those occupations into 15 business functions and 50 sub-functions across 16 industry sectors. This detailed mapping allowed the researchers to understand how applicable automation was to different economies, sectors and business functions. Then, leveraging work activity data, EY derived the amount of time that workers spent on individual tasks. Their analysis revealed that potential to automate tasks differs by more than 2X across sectors and up to 7X between functions. Functions are as varied as finance (heavily rules-based, where 80% of tasks hold potential for automation) and learning and development (with only 12% of work potentially subject to automation). Researchers found that every sector can transform roughly a third of its work.

Research conducted by the Institute for Spatial Economic Analysis (ISEA) noted the types of jobs that will be impacted by AI and result in job loss[3]. The data looks at urban centers, compares the job market, and shows trends of metro areas that will be affected. For example, according to the research, the Las Vegas-Henderson-Paradise, Nevada area has a 65.2% share of jobs that are automatable. This research, if predictive of the impact, will result in a re-engineering of many aspects of the workplace. In the public sector, and in many non-profits, administrative work often consumes funds that could be used for direct service provision. Evidence-based strategies to improve the impact of programs increasingly are being used to identify when and how programs are funded. The positive aspect is that algorithms and automation may be able to detect trends that would take many hours and people to identify. In the non-profit industry, this

[1] World Economic Forum, "The Future of Jobs Report, 2018"," p.8.
[2] "The future workplace: How to automate intelligently," EY, October 29, 2018.
[3] Institute for Spatial Economic Analysis, "Future job automation to hit hardest in low wage metropolitan areas like Las Vegas, Orlando and Riverside-San Bernardino," Jess Chen, May 3, 2017.

could be a silver lining because when funding cuts occur, they generally affect administrative overhead first.

Putting it into practice

As part of our research, we considered how organizations within the federal, state and local, and not-for-profit sectors are beginning to embrace AI. Selected actual, yet anonymized, examples are discussed below.

Introducing automation/ Robotics Process Automation (RPA): An organization within the US Federal government implemented RPA. To this organization, the advantages of RPA were evident. It streamlined routine processes to gain cost efficiencies and improve accuracy that would otherwise be compromised by human error. It also reduced the overtime hours and resulting stress that staff typically experience. In a way, RPA provides the staff needed without investing in additional Full Time Equivalents (FTEs). When RPA was first introduced to the organization, there was concern about replacement of jobs from personnel at all levels, from junior to seasoned senior-level employees. The organization's level of resistance peaked when employees first heard about the introduction of RPA. A majority of the anxiety felt was a result of employees' confusion about what RPA was and how it impacted their job security. To address these concerns, leaders held a series of briefings, town halls, and open discussions with staff to clarify what RPA was and what it was not. These public forum-style events were critical to mitigating resistance and encouraging acceptance among employees across all levels who now felt included in the decision to implement RPA. Over time, the staff became more comfortable with the idea of RPA as they learned what it was, what it would be used for, and how it would be implemented within their organization. During these sessions, leaders emphasized that RPA would improve rather than replace their jobs and highlighted that the time savings gained by RPA would be repurposed toward more strategic activities. This organization became more accepting of the idea of RPA more quickly due to their pre-existing familiarity with the concepts of Lean Six Sigma thanks to an ongoing internal initiative. Leaders were confident that RPA would improve the workforce and the organization's culture rather than inhibit it. Though change is not always welcome, this organization believes it can be exciting if improvements are clearly defined. For example, as soon as RPA training was offered to employees, staff not only signed up and attended, but they also volunteered to partake in additional training sessions. The organization has noted that key success factors include making every effort to secure stakeholder buy-in and identifying appropriate resources from the outset to successfully couple the implementation with any ongoing process improvement initiatives.

Improving mission effectiveness through AI in Public Safety: AI has significant impacts for the first responder community. One of the most important areas where AI can improve the mission effectiveness of first responders is around situational awareness. One example of this was demonstrated by a public data aggregation technology that could pinpoint the location of the shooter in the recent Mandalay Bay tragedy in twenty-two seconds. Social video feeds were analyzed to detect a muzzle flash that was a large enough anomaly to be pinpointed as the origin of the shots being fired into the music venue. Technology like this can work alongside public safety and response teams to improve their ability to identify and analyze signal information in

extremely short periods of time, and can prove pivotal to mission success. Similarly, public safety entities are examining gunshot pinpointing programs where a network of sensors spread throughout a city can allow public safety personnel to more quickly determine where shots are fired and integrate with existing police information to provide mug shots, warrants, and criminal history information before officers even arrive on the scene. AI can be a significant game changer for first responders. They currently carry and manage a myriad of items and tools that can be consolidated into a single smartphone or concentrated in centralized network communication centers. Thus, in the public safety realm, AI is seen as an augmentation to mission ability not a replacement of roles or personnel. This understanding mitigates much of the fear of job or role replacement risk which could hamper these endeavors.

Leveraging AI in a not-for-profit: A crisis hotline, founded in 2013 and headquartered in New York, provides services across the United States and in Canada. It receives thousands of texts and comprises a team that responds to people in crisis. An employee of the hotline may be monitoring 20 conversations of different counselors, with a great deal of complexity to the conversations. Using AI to assess the data with an algorithm, the team can benefit by picking up language that assists in predicting when a person is escalating in crisis in real time. The system uses coding to flag certain patterns. The tool has assisted in increasing the flagging of people in crisis and routing the calls. The system has also helped in the processing of the texts to the hotline. However, some people fear that the algorithms might be relied on too heavily and have unintended consequences, like incorporating the bias of the programmers or people providing the services. Institutional bias can occur with or without AI and it is one of the fears that surface when this topic is discussed. It seems like the advantages could outweigh the disadvantages, but protocols and systems need to be put in place to address these issues. Additionally, there must be a discussion of ethics and integrity around when and how AI is used for not-for-profits, especially for clients that have chronic or recurring behavior issues.

State and Local Transportation and AI: In examining the emerging impact of AI in local government, it seems to be highly concentrated in the area of transportation. From London, England, to Columbus, Ohio, to Dubuque, Iowa, connected vehicles and connected transportation infrastructure is emerging. We can learn lessons from the integration of Geographic Information Systems (GIS) into organizations. GIS offers a recent example of the challenges of integration. Years ago, a city in the US purchased GIS and created a position to manage, deploy and integrate a GIS system into the organization. With some notable exceptions like the Planning Department, individuals tended to look upon the GIS division as a resource for information, not as a tool integrated within their department with their employees directly using that GIS tool. However, the city has now created a new position that will allow the GIS division to assist departments with training and other services to integrate the GIS tools within city departments. It will also bring in not-for-profit partners to help them integrate GIS tools within their organizations to leverage the power of this big data source and lead to better decisions, more powerful cooperation among partners, improved measurable outcomes and most importantly, more impactful outcomes for customers. This approach is further supported by the implementation of work teams to enhance support and sharing across departments and

organizations. AI, matched with GIS, has tremendous potential to improve the customer experience.

Moving forward: a road map for success

Overcoming the challenges and fully making AI work within an organization requires detailed planning and creative solutions. Practical strategies include operational, human resource, change management, process and policy considerations. We offer the following short list of practical suggestions to assist with the AI journey.

Strategy and planning: Clearly defining the program's purpose and articulating the benefits thereof to your teams are essential first steps.

- *Align AI to your mission and purpose*: Government and not-for-profit organizations are intrinsically mission- and purpose-driven. Explain to your employees how AI can be used to free them up from work that is not directly related to their mission and allow them to focus on higher level work that is more impactful. A clear purpose statement is powerful because it shifts an individual's focus from the tasks they perform to how all of the work done by a unit contributes toward a larger goal. Work cultures that focus on culture are ready and able to have work redefined by automation without morale loss, organizational anxiety or talent flight[4].
- *Be purposeful about the implementation of AI:* It is not just about implementing technology for technology's sake. Help your teams to understand this and stress collaboration. EY notes that it is important to "connect technology and human capability in a digital ecosystem". It really is about "collaboration - human and machine combining forces to deliver the best that innovation offers." [5]
- *Define the scope*: Some processes and functions are open to AI – others are not. Carefully assess those that could be candidates for AI deployment and be ready to admit when a process is not viable.
- *Define and describe the future state of the business function*: Consider the blend of the human and automated workforce. Consider roles, responsibilities, and how the two will interact, and describe that interaction clearly.
- *Be aware of algorithm bias*: As you plan, be aware that an algorithm bias may exist. It is important in your planning processes to include quality checks early in the process. For example, an algorithm bias receiving a tremendous amount of attention is facial recognition software that does not recognize people of different ethnic backgrounds. It is important to have this issue highlighted early in the process given the multitude of stakeholders that government and not-for-profit entities both engage with and employ.

[4] "The future workplace: How to automate intelligently," EY, October 29, 2018, New York.
[5] "Is your digital strategy built on human and machine, or human with machine?" EY, April 2018, New York.

Alignment and Governance: Considerable change management will be required. Take the time to identify your stakeholders and your change network. Carefully identify the executive sponsor and make sure they are prepared to communicate, sponsor and model the change.

- Build a structured communications plan that articulates the strategic benefits and the operational implications, and answer the "What's in It for Me" question at the beginning of the process.
- Build your talent pool because you may not be able to rely entirely on your current workforce. Work with your HR and recruiting teams to identify talent both within and outside of your organization that is AI ready, willing and able.
- Assess the readiness of your workforce and AI's impact: What is their current appetite for AI? What is their functional level of capability? What skills do they have and what would be needed to close the gaps? Who might be displaced? What training or transition plans do you have for those whose jobs are eliminated or significantly structured?
- Make sure you have the appropriate policies in place and that your employees are aware of them. For example, AI will introduce potential privacy considerations. Have a policy around this and make sure it is communicated to the organization.

Process development and testing: As you move through your AI journey, it is essential to develop, and potentially improve, your current processes to prepare them for AI deployment.

- There will be several rounds of testing (user acceptance testing being a key event) once you have automated your selected processes.
- Testing and process improvements are excellent ways to engage your workforce in the change, so that this becomes something that they are actively engaged in and not something that is done to them. It helps to build commitment to AI, not just compliance.
- Create a pilot project and identify people to lead it. For this trial run to be successful, leaders should outline what each team member will be doing and train them beforehand.

Ongoing management: AI implementation is not a "one and done." Success will come from the persistent focus on continued buy-in of the organization:

- Once you have implemented your AI, make sure that your workforce is bought into and aware of the process for maintaining it.
- Maintenance activities may include scheduling of bots, real-time monitoring, communication with application owners, and infrastructure improvements.

Concluding thoughts

The fundamental premise of this paper is that AI does not necessarily replace or eliminate jobs. Rather it can enhance and create opportunities for new jobs and better opportunities for workers. Understandably workers are concerned about AI eliminating the need for their position. However, with proper application, AI can be a strategy for improving the services of an organization and reaching higher levels of quality. Departments and organizations may be better positioned to serve their customers with the amount of time employees have available to

complete their responsibilities. In 1981, the first personal computer was introduced and in 2007 the iPhone emerged. Government has integrated these tools into their operations, as they will with future technology improvements. Based on past experience with the introduction of information technology tools, AI will change what an employee does, or how they do it, but it will not eliminate all positions. If introduced in the right way, AI can enable organizations to achieve more, and produce better and timelier outcomes for the citizens being served.

Karen Shrum *is a principal in the Government and Public-Sector practice of Ernst & Young LLP (EY). She has 20 years of experience leading government and private sector organizations through transformation. She advises government organizations on their complex transformational and management challenges. Ms. Shrum formerly served with the Department of Treasury, Office of the Comptroller of the Currency, where she delivered complex multi-million-dollar IT related programs. She holds a Masters of Business Administration (MBA) in Organizational Behavior from the George Washington University and is a Project Management Professional (PMP).*

Lisa Gordon *is the President and Chief Executive Officer of Atlanta Habitat for Humanity – one of the top 10 Habitat for Humanity International affiliates in the United States. She is a recognized leader in transformational redevelopment efforts for quality affordable housing and neighborhood revitalization. Ms. Gordon joined Atlanta Habitat in July 2015 and set the nonprofit homebuilder on a new course to become a catalyst for holistic neighborhood revitalization. As a leader in urban redevelopment and government service before joining Atlanta Habitat, she was Vice President and Chief Operating Officer for the Atlanta BeltLine, Inc., cabinet member in former Atlanta Mayor Shirley Franklin's administration and has held City Manager and Assistant City Manager positions. Ms. Gordon is a Fellow of the National Academy of Public Administration*

Other contributors include: Shane Sowards (Senior Manager, EY) and Peter Mangan (Senior Consultant, EY).

A Public Administrator's Practical Guide to Ethics and Artificial Intelligence

Prepared by Priscilla M. Regan, Subgroup on Ethics and Bias Chair; and Karl Maschino

The goal of this working paper is to add to the ongoing conversation about the role of, and ethical issues raised by, the use of artificial intelligence (AI) in public administration (PA). The intent is not to repeat or even synthesize what has already been written in other places. Appendix C provides an annotated bibliography of several relevant articles and reports; however, this list should not be considered exhaustive of the literature. After reviewing relevant documentation, we draw upon our experience as scholars and practitioners in public administration and public policy to help identify the ethical issues raised by the use of AI in public administration in a way that will be of value to the PA community. This analysis focuses on the civilian side of government, including law enforcement, education, health and social services, and does not address the issues raised in the defense and intelligence sectors.

Overall Context

The fundamental aspect of AI that raises ethical issues is the fact that there is not a clear explanation for how and why an AI application reached its conclusion – be that conclusion a decision about who should receive what benefits or punishments, or what geographic areas will be most vulnerable to storms or forest fires in 25 years, or how effective a particular drug is likely to be for a group of patients with certain characteristics in common. The opacity of AI systems in any area of PA challenges the traditional responsibilities of administrators regardless of substantive policy area or whether decisions directly involve citizens. The core issue that is raised involves the ***accountability*** of the AI system and the ***responsibility*** of the administrator who manages the program in which AI is used. Therefore, when considering use of AI systems, an administrator should consider:

- In what ways is the deployment of an AI system likely to change my ability to oversee the program? What am I likely to understand and what might I not be able to understand regarding a decision or conclusion that an AI system reached?
- When an action or "decision" is made that is questioned by supervisors or by Congress, who is responsible? The Program Manager? The contractor who wrote the code? The AI itself as it developed using machine learning?
- How do you consider the role of judgement, which is the key value added proposition of senior administrators?
- What is the acceptable level of autonomy in AI? Will this vary depending on the application? What factors will determine the acceptable level of AI autonomy?

- Will explainable AI be enabled to testify in front of Congress? For an explanation from DARPA see (https://www.darpa.mil/program/explainable-artificial-intelligence) of explainable AI, an area on which they are working.

In answering these questions, indeed in answering all the ethical questions we identify, the critical step is understanding what is agreed to in the ***contract*** and ***terms of service*** that is negotiated and/or accepted at the time the AI system is employed. In most instances, the program or agency will likely be procuring an AI system or application from an outside vendor. If the AI product is being developed by agency personnel, then then the answers to such questions should be addressed in developing the work plan or specifications for the AI system. This is the point at which the relationship with the AI product is defined, the role of AI in agency programs is specified, and the responsibilities of the AI vendor or AI developers are established. Additionally, the contract or terms of service should address the question of who owns information created by AI, the government agency or the contracted entity performing the function, as well as who owns the knowledge created by the application of AI.

The control lever of contracting is critical; thus, program managers and procurement officials need to work together with data scientists and knowledgeable technical staff to ensure that contracts address key questions of responsibility and accountability. In determining the answers to these questions, program managers, technical staff, and procurement officials may well need to employ standard ***risk analyses*** and document their findings, including their reliability and validity.

Specific Ethical Issues

Information Privacy

In many areas of administration, e.g. health care, social service programs, education and labor, AI systems rely on the incorporation and analysis of personal information that was collected for other purposes. Under the Fair Information Principles, codified in the Privacy Act of 1974, this would arguably be considered a "secondary use" that was not consistent with or considered when the information was originally collected. In some instances, agencies might argue that the use of previously collected information for a somewhat related purpose fell under the "routine use" exemption in the Privacy Act and other agencies might regard a Privacy Impact Assessment as necessary before implementing an AI program that employed personal information. But agencies appear to have discretion here under existing policies.

Moreover, there is now a treasure trove of personal information available to agency personnel and AI vendors that might be considered relevant or useful in developing an AI system. The possibilities of using such information raise the ethical issues identified in discussions of "big data." Three government reports by the FTC, PCAST and the White House addressed these issues, as well as several reports by outside groups (see Appendix C). With big data, there is more collection of information, by more parties, about more aspects of an individual's life, and with more granularity about that life. But the issue here is not merely the amount and detail of information or even the qualitative changes that quantity does not convey. The issue is also how

much of big data collection takes place without the individual's awareness. Moreover, enhanced digital storage capacity combined with improvements in computational power and developments of more sophisticated algorithms for analyzing data have enabled organizations to probe and dissect datasets to extract potentially useful information. The entire enterprise of big data challenges previous ideas about how to limit data collection about individuals and how to involve the individual in the process of data collection and subsequent uses so that the individual could exercise some meaningful control.

Public administrators should be aware that the uses of existing agency databases of personal information and/or the culling of personal information from other sources, such as public records or social media postings, would raise ethical questions. Some specific questions that administrators might ask include:

- What personal information is being used in the AI systems?
- What are the sources of that information? How reliable are those sources? What are the risks that inaccurate, incomplete or irrelevant information will be used in the AI system?
- Are individuals likely to perceive the use of this personal information in the development of the AI system as appropriate and reasonable or as "creepy" and improper?
- Are the inferences that will be made based on the incorporation of non-agency datasets consistent with previous agency procedures and norms regarding how decisions are reached?

Anonymity

As more social relationships and practices are rendered as data points, it is more difficult for individuals to maintain the "practical obscurity" that existed somewhat naturally in the past. Searches of large datasets, using sophisticated computer tools, rather quickly eradicate what had been high transaction costs on finding meaningful information. The distinction between personally identifiable and non-identifiable information has largely evaporated. Current practices for anonymization or de-identification of information about individuals may not be effective as big data make re-identifying data far easier. As the research of Latanya Sweeney (see Appendix C) and others has revealed, few characteristics are needed to uniquely identify an individual, which makes it more difficult to anonymize databases by removing some characteristics, because the remaining characteristics are likely to be sufficient to identify individuals once a database is merged with other databases and searched using sophisticated algorithms. Although computer and data scientists continue to develop more effective tools for de-identifying data, more effective techniques may render the information less useful. This is particularly problematic in areas like health data.

Public administrators should be aware of these issues involving anonymization, de-identification and re-identification and might ask the following questions:

- How should we calculate the risks that someone might be re-identified?
- What data fields should we delete from the algorithm?

- How often should we audit what the AI system is producing to determine whether individuals may be able to be identified?
- Are there certain types of data that we should not include at all, e.g. date of birth, address or health status, in the development of the AI system?
- What are the benefits of being able to use data sets that potentially can be re-identified but that may yield socially or individually important insights?

Discrimination

The question of whether AI systems may contain biases is arguably the largest ethical issue and the one that has received the most attention. One of the cornerstones of administrative decision-making is the traditional due process requirement of ensuring fairness and consistency. Similarly, situated individuals should be treated similarly and there should be no discrimination based on certain characteristics such as race, ethnicity, religion, gender or other personal attributes – or based on factors of which they are not aware. Even if AI systems do not use these characteristics in their algorithm, other characteristics and data may serve as proxies – very possibly without the knowledge of the developers of the AI system or the administrators using the system. Additionally, as the AI system learns and modifies itself, it may develop a proxy for a protected characteristic that was not in the original algorithm. Finally, the AI system may sort people in new ways based on a combination of characteristics and discriminate on these bases in ways that are possibly socially unacceptable.

One concern that has been raised repeatedly is that the data scientists and computer programmers who develop AI systems may approach the system in terms of efficiency and formalization. As a result, they may fail to take into account the broader programmatic context in which the AI system will be used. Having a team of administrators with a range of backgrounds and expertise participate in and advise on the development of AI systems may help to address this concern.

Given the numerous ways in which administrative programs categorize individuals and render decisions with often grave implications for individuals, any indication that an AI system may be discriminating among individuals needs to be taken seriously. Questions of possible bias and discrimination in AI systems are particularly critical in areas of administration where individuals may be awarded or denied a benefit, a consequence such as a sentence in a criminal case, access to a service or treatment, a grant or waiver, etc. At this point, the evidence indicates that existing AI systems often yield discriminatory outcomes, particularly with regard to race.

In determining whether AI systems discriminate among individuals in unanticipated or socially unacceptable ways and how to best deal with potential or real bias in AI systems, public administrators might ask the following questions:

- How can AI systems be tested before they are employed to ensure that they will not discriminate among individuals in ways that have traditionally been prohibited or to determine if they are discriminating among individuals in unanticipated ways?
- How often should AI systems be audited once they are employed to determine if they are inserting unexpected biases into decisions?

- What information should be provided to individuals who are subject to decisions reached by AI systems?
- In developing or purchasing AI systems that will be used in decision-making about individuals, what experiences and expertise should be involved in developing and evaluating the AI system before it is used?
- What redress or grievance procedures should be available to individuals who believe they have been unfairly treated as a result of an AI system?

Autonomy

One of the newer tools of government being used in administrative programs is "nudging" people to more socially beneficial choices. Some ethicists question whether such programs are manipulative and restrict the space for human judgment, therefore representing a problematic, if not an unacceptable, role for government. One prominent advocate of nudging, Cass Sunstein (see Appendix D) argues that these ethical issues should be evaluated on the basis of whether nudges promote or undermine welfare, autonomy and dignity. He views nudges as generally helpful and defensible on ethical grounds as they often promote social welfare, provided that they incorporate transparency and accountability as safeguards. He cautions, however, that to be ethical nudges should steer people in certain socially valued directions but at the same time allow them to make their own choices.

Nudging messages are often used in contexts where administrators are trying to guide people to what program research and evaluation consider to be more appropriate behavior without requiring that behavior, e.g. healthier eating habits or exercise regimes, saving for retirement, recycling for a better environment, etc. They are often seen as a form of education and information dissemination. Such messages evolve from simple education to nudging when they incorporate insights from behavioral economics, marketing research, and/or psychology about what is likely to motivate people generally or motivate a particular group of people.

Ethical questions regarding the appropriateness of nudging are especially acute if the nudges are targeted to particular individuals as a result of an AI system. With AI systems, a group of individuals can be provided certain information as a result of being identified by an AI system and "nudged" to behave in a certain way or to believe certain things while other individuals are either not "nudged" or are "nudged" in a different direction – raising not only issues of individual autonomy but also issues of discrimination as discussed above. The Facebook/Cambridge Analytica incident from the 2016 election and the targeting of particular messages to segments of the population based on inferences about their political sensitivities illustrates the ethical problems that might be raised.

If public administrators are considering the use of an AI system to target "nudging" messages to individuals, some of the questions that should be considered include:

- Are messages suggesting options to individuals or leading them to a predetermined choice?
- Can you identify a line between influencing people and manipulating them?
- What is the timing of the nudge in terms of when an individual is likely to need to make a choice?
- How susceptible is the targeted population likely to be to a nudge?
- How does the AI system design a message and/or select a targeted population? How different is this from what program administrators would do without the assistance of the AI system?

Surveillance

One of the capabilities of AI systems that is particularly attractive in public administration is their ability to search vast quantities of data and identify patterns. This capability coincides with the fact that individuals are not only now revealing, especially on social media sites, more information about their activities, preferences and relationships but also now are emitting things about themselves that they do not even fully realize (Internet of Things). The sophisticated prediction in AI systems transforms the ability to track and analyze such data in ways that have not yet been identified and understood.

The amount of data that is being generated, as well as the ease and low cost with which various data streams can be surveilled, make such passive surveillance and the use of the data in AI systems very attractive. There are however questions about the relevance of such data to the program for which it is collected. There is often a tendency to collect as much data as possible on the chance that it might be helpful, reveal new information, or be relevant in some unanticipated way. Information overload can be a consequence, as might confusion as to how to regard the reliability and relevance data that is collected. Too much information may simply slow administrative decisions. Moreover, there are ethical questions about the appropriateness of surveilling certain types of data, e.g. social media sites. Finally, there are questions about whether the AI system is using the proper techniques or meaningful fields in its search and analysis of the data streams.

In considering whether to use an AI system to surveil data streams that are not originally collected by the agency, public administrators might consider the following questions:

- What kind of data has the agency used in the past in making these decisions? Is the use of these databases and AI systems to search these databases consistent with past practices? What value is to be gained?
- How reliable, relevant and accurate are the databases that your AI system is monitoring? If it is an appropriate database given the mission of the agency and programmatic goals, is the AI system extracting the appropriate information?
- Are individuals whose data are being monitored likely to regard the surveillance and extraction of data as legitimate given the previous practices of the agency or to regard it as "creepy?"

Priscilla Regan *is a Professor in the Schar School of Policy and Government at George Mason University. Prior to joining that faculty in 1989, she was a Senior Analyst in the Congressional Office of Technology Assessment and an Assistant Professor of Politics and Government at the University of Puget Sound. From 2005 to 2007, she served as a Program Officer for the Science, Technology and Society Program at the National Science Foundation. Since the mid-1970s, Dr. Regan's primary research interests have focused on both the analysis of the social, policy, and legal implications of organizational use of new information and communications technologies, and also on the emergence and implementation of electronic government initiatives by federal agencies. She is currently a co-investigator on a Social Sciences and Humanities Research Council of Canada's eQuality grant exploring big data, discrimination, and youth. Dr. Regan is a Fellow of the National Academy of Public Administration.*

Karl Maschino *is the Chief Administrative Officer and Chief Financial Officer for the U.S. Government Accountability Office (GAO). He is responsible for the internal operations of GAO, including information technology, finance, budget, human capital, training, labor relations, facilities, security, and field operations for an agency of approximately 3,000 staff, with locations in Washington, D.C., and in 11 field offices around the country. Previously, Mr. Maschino spent 16 years with the U.S. Department of Justice, most recently serving as the Executive Officer of the Criminal Division. In that role, he was responsible for the operational support of a $600 million organization of over 1,000 staff located throughout the United States and in over 60 countries around the world. Prior to entering public service, he worked in the fields of sustainable economic development, and international higher education in university, city government, and nonprofit settings. Mr. Maschino is a Fellow of the National Academy of Public Administration.*

Artificial Intelligence and the Teaching of Public Policy and Administration

Prepared by Dr. Alan R. Shark and Anders Shropshire

Introduction

While much on-going discussion has focused on how information technology can be integrated and introduced within the PA curriculum or offered as a separate intro course, artificial intelligence itself has accelerated the need to turn thought into action.

Applications of AI in the form of machine learning have already entered the public management domain. It has taken many forms, including the use of chatbots in citizen engagement systems, as well as the integration of smart, language-based, interactive systems that field inquiries through such recognized systems as Siri or Alexa. Today, autonomous transport vehicles are teaching systems to learn about real-time road conditions including distinguishing between objects, people, pets, and unexpected scenarios. Other government AI applications could be described as "augmented decision-making." Augmented decision-making is used for cyber security monitoring, public policy modeling, database anomalies seeking discrepancies in prescription drug duplications, and other forms of waste and abuse. AI is now being used in the field of public safety in areas such predictive crime analysis and prevention. AI systems are becoming essential tools in the PA practitioner toolbox where speed, efficiency and effectiveness can save huge amounts of time and money, and reach conclusions faster than humans could even imagine.

AI, while gaining much attention these days, is not as new as some might believe. The field of AI goes back at least 25 years, but because of a confluence of technologies, it is indeed expanding and is now reaching the halls of government. Like the human brain, machine learning not only mimics how we process information and experiences, but it learns from them as well. Add robotics to machine learning, and you have devices that can easily replace human jobs that are dependent on repetition. Some economists have predicted a major upheaval in the workforce by the year 2050. Others claim that new jobs will be created to make up for most of any loss. Much of AI's growth can be attributed to:

1. Advancements in complex algorithms.
2. Dramatic increase in speed and computing power.
3. Ability to digest data from various sources (voice, videos, text, social media, etc.)
4. Ability to store and retrieve massive amounts of data in fractions of a second.

With every advancement in AI, governments at all levels will continue to face many new challenges, including ethics, privacy, human control, policy bias, predictive analytics, decision-making, citizen engagement, planning, and the future of work.

Until recently, AI has been taught primarily in schools specializing in computer sciences, engineering, robotics, and computer programming. Those who desire to enter public service and choose a degree in either public policy or administration are mostly left out of any technology focus, let alone AI instruction.

Therefore, central to the discussion of *AI and the Teaching of Public Policy and Administration* is what role, if any, should or can be played by schools of public affairs and administration? Rather than take a prescriptive approach, this paper begins with a simple but important survey aimed at better understanding the academic environment and perceptions of key academic stakeholders regarding AI.

We offer a model for how AI could be introduced as a stand-alone course or individual or group project in Appendix A.

Results of Academy Survey of Public Affairs Schools/MPA Programs Teaching Artificial Intelligence Topics

Survey Methodology

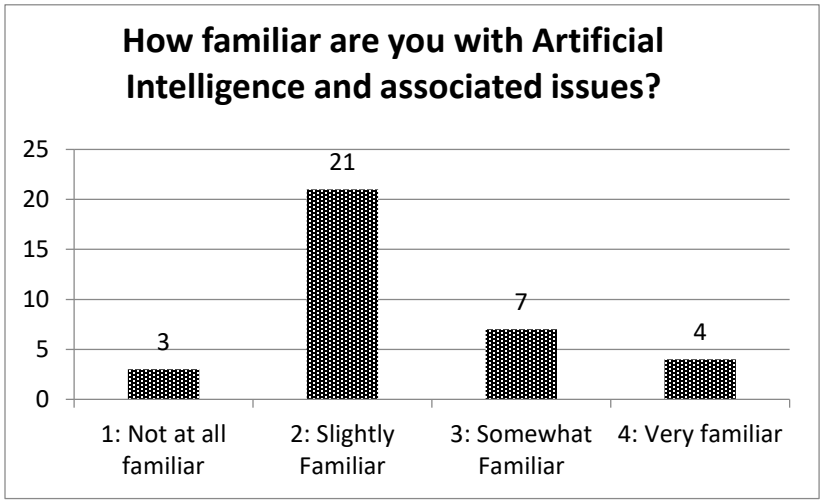
As a part of its research on the role of AI being used in Public Administration, and issues associated with that role, the Working Group on Artificial Intelligence was interested in what response schools of public affairs were having to the emergence of AI in the public sector. A survey was sent out to the deans or MPA program coordinators of public affairs schools which probed them on the question of how AI fit into their school's curriculum and how they should adapt to this new issue. It is the opinion of this working group that incoming public administrators should have some awareness of these issues, as they are increasingly likely to encounter them at some point in their careers. There is a growing list of instances, many controversial, where AI is deployed in public sector settings. Additionally, the emphasis placed on workforce reskilling in the President's Management Agenda, the administration's effort to modernize government for the 21st century, further indicates the need to retool curriculums as data, AI, and advanced computing become ever more central to government work. Schools of public affairs are on the frontline in training the next generations of public administrators at all levels of government. It is important that they become familiar with the tools of the future – which in more and more cases are proving to be the tools of today.

The informal survey set out to understand what place AI has, and should have, in current curriculum according to the program directors of public affairs schools. A list of all operating public affairs schools was taken from the NASPAA (Network of Schools of Public Policy, Affairs, and Administration; a trade association offering program accreditation) website and narrowed only to include schools in the 50 U. S. states and the District of Columbia. Contacts from each school were identified and twice emailed the survey, approximately one week apart. In all, 262 schools were contacted. The survey received 35 responses for an overall response rate of 13.4 percent. The responses were collected between July 18th and August 1st 2018. The results of the survey are presented below.

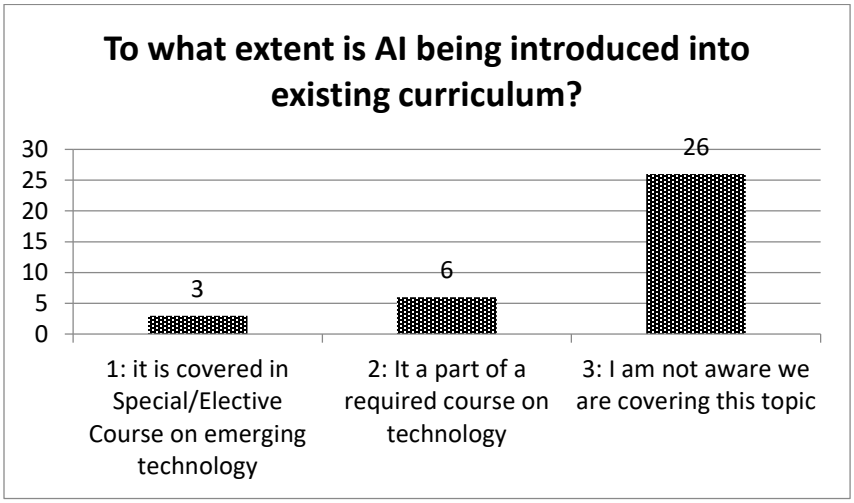
Presentation of Results

The first question asked the respondents to describe how much they knew about AI technologies and topics on a 4-point scale. As can be seen, most respondents have little exposure to these topics, with just 30% describing themselves as "somewhat" or "very" familiar. This indicates that much of the sample have not been exposed to the breadth and depth of the AI literature. This may influence how important they feel this issue is, particularly in the realm of public administration. Respondents who are 'slightly' familiar may be unaware of the many government roles AI is

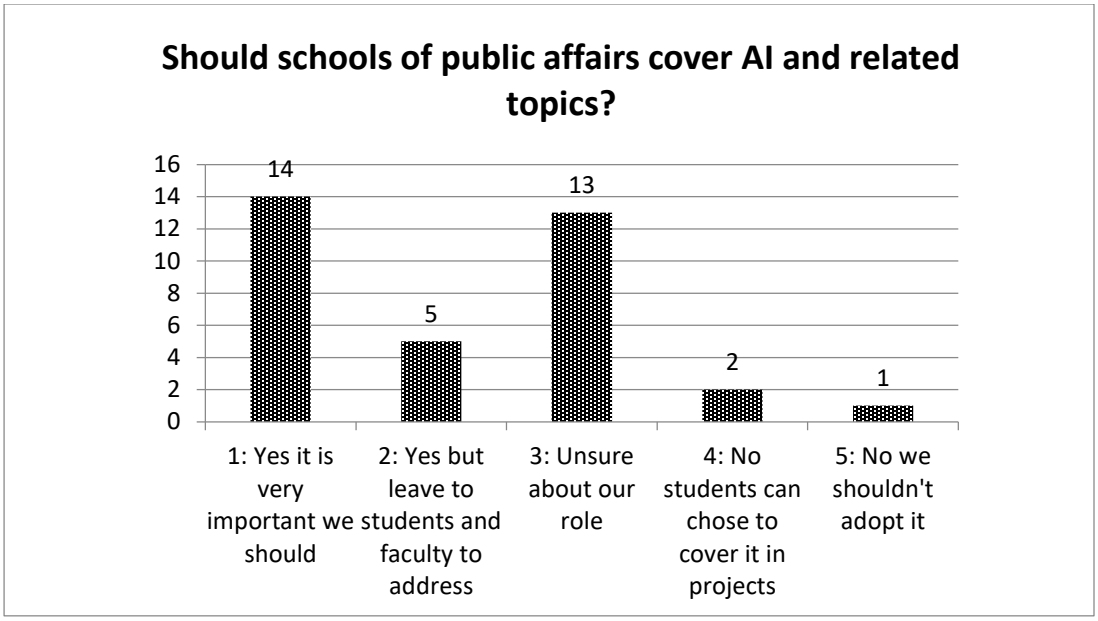
filling across the country. They may also not have known of the concerning ethical issues relating to privacy, bias and discrimination, and more that are being debated in AI literature. Finally, they may not grasp the incredible potential these technologies have to improve and revolutionize government. This finding indicates that AI is not yet on the radar of most trainers of future public administrators.



The next question asks whether the topic is being covered in schools already as a part of a required or elective course on emerging technologies. 75% of programs indicate they do not cover this topic, echoing the results of the previous question.

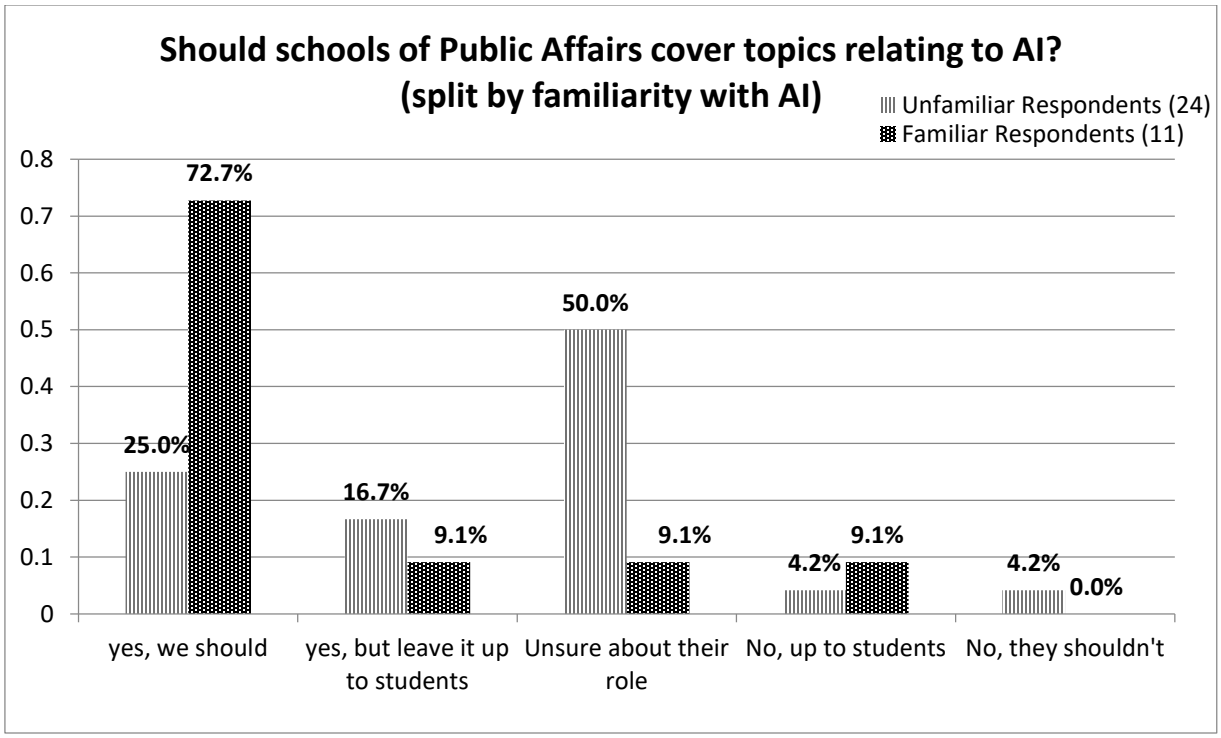


Next, the questions began to examine what position AI should occupy in these schools' curriculums. Respondents were asked to report whether they believed AI should be adopted into existing curriculum, either as a part of a new or existing course that could be an elective or required. The results indicate dueling feelings of urgency and skepticism. 40% feel that these topics are very important and should be adopted, with an additional 15% feeling it is important, but not up to the administration to decide. Only 8% of respondents reject the idea outright. Still, another 37% are cautiously uncertain about the role of their school.

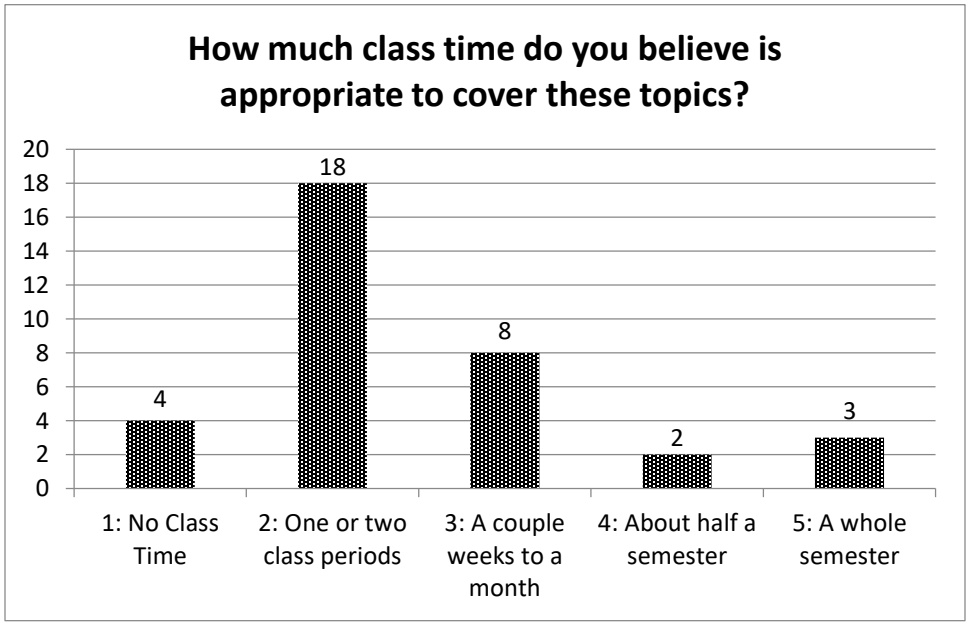


The cause of these competing sentiments is revealed upon more careful analysis; the respondent's self-reported knowledge of AI heavily relates to how essential they find it. To perform this analysis, those who identified themselves as "Not at all" or only "slightly" familiar with AI are condensed into the 'unfamiliar' group, while the "somewhat" and "very" familiar respondents formed the 'familiar' group. Twenty-four respondents, or 68%, were in the unfamiliar group, with the remaining eleven in the familiar grouping.

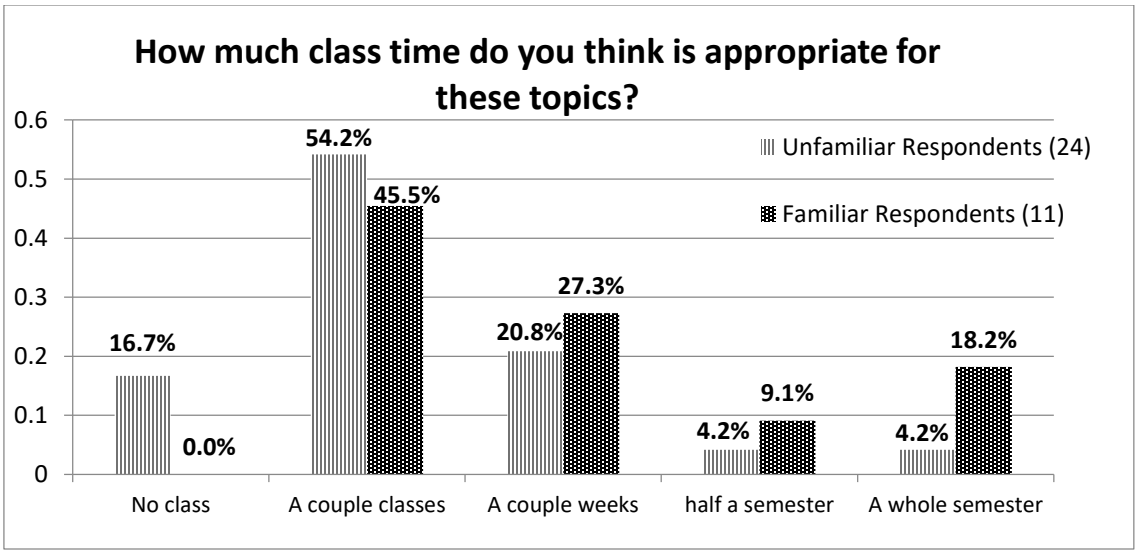
Comparing the distribution by groups indicates that the familiar group identified this as an important issue at much higher rates than the unfamiliar group. A full 72% of them feel that the curriculum should adopt AI topics. On the flip side, of the thirteen respondents who were unsure about their role, 12 came from the unfamiliar group. Of the fourteen that said it should be adopted, eight came from the familiar group. Therefore, it is possible that after learning more about AI's growing role in government and other organizations, respondents feel it is more important to incorporate into the curriculum.



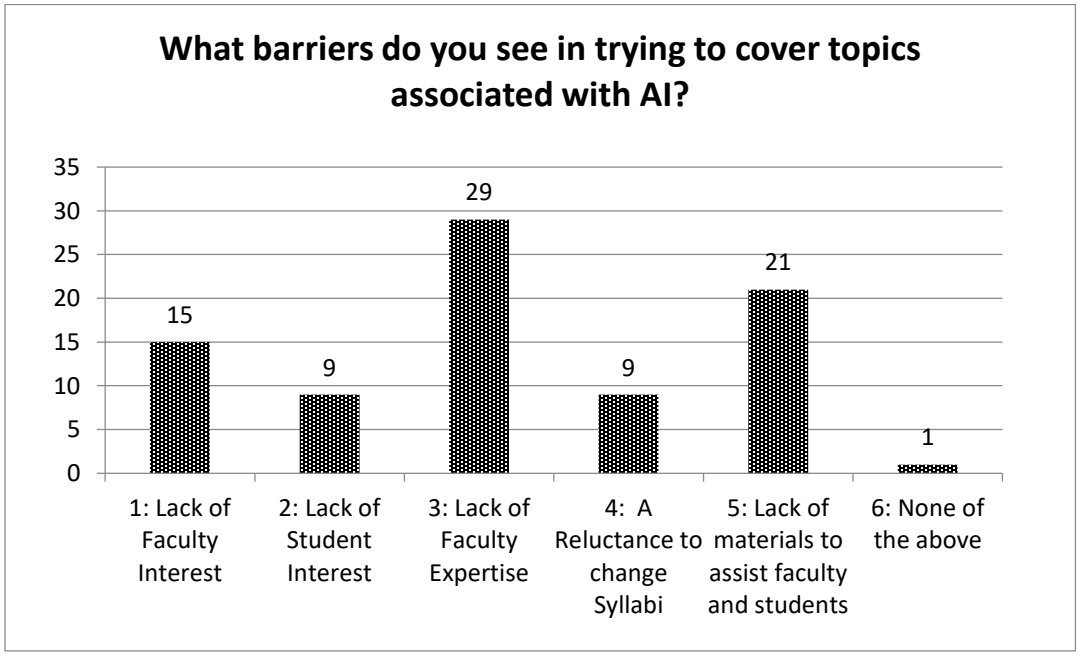
The next question asks the respondents to speculate on how much time would be needed to cover these topics, given the limited space school curriculums have. In general, the respondents do not feel that too much class time would be needed, with the majority indicating that one or two class periods would be enough. A quarter believe that two weeks to a month would be more appropriate, but only 8% feel a whole semester long course is necessary.



The responses here vary by familiarity with the topic as well. Those more familiar with the topic tended to advocate for more course time, though the difference is not as stark as in the previous question.



The last question asks respondents to identify major barriers that are preventing or would prevent the school from instituting a class or unit on AI topics. Respondents were asked to check all that apply for this question, and an 'other' option was offered in a subsequent question. The biggest barrier clearly is a lack of faculty expertise. Over 80% of respondents identified this barrier, which matches the lack of familiarity with the topic identified in the first question. 60% of respondents see a lack of materials as a barrier. It is likely that these both relate to the lack of coverage of AI in most academic journals and the emergent state of the technologies. A lack of interest among faculty and students appears to be a barrier for some, but the majority does not share that sentiment. One respondent noted the fact that "NASPAA Standards do not call for information and technology as a required competency" may be a barrier for some programs. Overall, it seems that the unfamiliarity with this topic is the biggest barrier.



Discussion

These results indicate that the leaders of PA schools and MPA programs are hesitant, but open to adopting AI into their curriculum. A recurrent theme in the results of this brief survey is a lack of familiarity or expertise on this issue among public administration academics. Most have simply had little exposure to this topic, which is understandable given that AI has only begun to emerge at a rapid pace and that it is often associated with the computer science field. An important finding is that those who self-identified as being more up to date on AI topics found it more important to cover the issue. They see the growing role of these technologies in public administration and the necessity of training the next generations of students in that field.

Most respondents indicate they feel that at least a couple of classes could be spent focusing on AI. One respondent suggested it could fit in a survey course. Another respondent mentioned a course currently under development in their school entitled, "Artificial Intelligence and Machine Learning", and another remarked that they offer elective courses on the subject but would like to integrate it into the required courses in the future. But most respondents worried about a lack of faculty expertise and materials hindering their efforts to teach these new and complicated subjects. As AI grows, the field should become more informed and these concerns should go away. In the meantime, The National Academy of Public Administration is assembling materials to help introduce faculty to these topics and create draft syllabi (see Appendix A). The time seems right to begin piloting these topics in the classroom.

Conclusions

Artificial Intelligence will change public administration across the board. Over the next decade, lawmakers and senior public managers will be making monumentally important decisions about the role of AI in society while pioneering the introduction of AI in state, local, and national government. It is imperative that we provide them the tools to do so knowledgably, and not leave them dependent on the advice of computer science professionals. Brauneis and Goodman[6] effectively convey the dangers of letting inexperienced and naïve professionals make these important decisions. The literature shows that there are immense unanswered questions about ethics, bias, privacy, and legal implications of these technologies and AI researchers have called for interdisciplinary discussion on these topics. Public administrators must understand both the capabilities and limitations regarding AI, how and why they could go wrong, and what to consider when deploying them if we are to gain their benefits. It is time for the public policy and administration field to start paying more attention and to start preparing the new cohorts of students to address these questions. It is increasingly difficult to imagine students graduating with either a MPP or MPA and not having a healthy and critical understanding of the very tools that can improve public management at its core. This vital area should not be left to scientists, programmers, and engineers alone.

[6] https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3012499

Dr. Alan R. Shark *is the Executive Director and CEO of Public Technology Institute (PTI) now a subsidiary of CompTIA. Dr. Shark's career has spanned over 30 years as a highly recognized leader in both the nonprofit management and technology fields, with an emphasis on technology applications for business and government. He is an as associate professor at George Mason University's Schar School of Policy and Government. Dr. Shark is a Fellow of the National Academy of Public Administration.*

Anders Shropshire *is a senior at the University of Wisconsin studying Political Science and Sociology, and a first year Masters of International Public Affairs student at the La Follette School of Public Affairs. He interned with the Academy during the summer of 2018 and served as the AI Working Group's principal researcher and, in many cases, its principal writer.*

Appendix A: AI 101 (Model) for Public Policy and Administration School

Introduction: What this course is about

Recent advances in Artificial Intelligence (AI) hold the promise to revolutionize the way in which human affairs are conducted, including in the world of government and public administration. Its power to perform monotonous and repetitive tasks, predict, and analyze give it incredible opportunity to streamline government and fundamentally change civil service. But for all the promise AI holds, it is often fundamentally misunderstood and comes with many important challenges, limitations, and ethical and governance questions public administrators will have to address. Before many agencies can even consider deploying AI, ancient legacy IT, scattered and incomplete data, and a lack of expertise are just a few challenges to be overcome. This course will be designed to prepare future public administrators to identify uses for AI within government and responsibly and thoughtfully implement it into their operations. The course will begin by defining AI and outlining the many different technologies that are often associated with AI and overviewing what applications they hold in public administration. The course will then explain the greatest internal challenges of the technologies before moving on to challenges within government that pose barriers to implementation. Lastly, the class will close by helping develop a framework for how future public administrators can identify internal administrative or external service issues to be resolved by AI and develop proper and effective systems to successfully implement AI technologies.

Unit 1: What is AI? A brief survey

- What is AI? Artificial vs Augmented Intelligence, AI as statistical analysis and modeling, AI as pattern recognition, AI Expert Systems, What AI can't do, Where AI is heading
- Technologies: Machine Learning, Facial Recognition, Chatbots/text to speech, etc.
- Training: How to teach AI – machine learning/data sets vs expert systems; supervised, unsupervised, and reinforcement learning
- Robotics

Unit 2: How to use AI: A survey of existing or possible practices in Public Administration

- Customer service reps, outsourcing paperwork, reviewing applications, helping with urban planning, surveillance, predictive policing, building personal profiles, combing through massive data sets like tax records, cyber security, becoming legal experts or experts in other processes, and more

Unit 3: Reimagining Public Administration after AI

- What the AI will do and what it means for civil servants. How do we begin to change the model of what it means to work in government? Do we downsize government? What do we do with money saved? Do we risk losing personal autonomy, institutional expertise, or decision making ability over time with AI? What new skills will we need?

Unit 4: Issues in AI and how to resolve them. Touching upon problems of bias, transparency, privacy, and other concerns.

- How biased data affects machine learning; how to screen for biases, both evident and subtle (i.e.: we have policing data, not crime data what does that mean?)
- The Black Box: how AI makes its decision and how we keep it open in a government context
- Security risks: new types of attacks and vulnerabilities – adversarial examples Ethics: Consider the dilemma of letting computers make human decisions
- Privacy: How to properly handle personal data that may be sensitive. With the ability for AI to create complex profiles of citizens from a diverse range of data, we might consider certain protections to implement to prevent inappropriate use by authorities
- Regulations: how to regulate a rapidly developing technology without stifling innovation
- Legal Liability: Assigning fault for the damage caused by an AI error. Examine the role of developers, trainers, operators, etc.
- Emerging case law: regularly updated to capture new cases

Unit 5: IT Infrastructure Requirements: How a new head of an agency would need to adapt to implement it.

- Hardware: computers, data storage
- Software: Data and data sharing,
- Human Resource: IT professionals, data scientists
- How to promote these advancements? Technology Modernization Fund

Unit 6: Project Development: How a public agency can implement a successful AI program.

- Setting goals to implement AI in ways that make government more human and human centered
- Maintaining Human Oversight
- Identifying other sources
- Implementation Stages: Assessment, development, training, evaluation, deployment

Unit 7: Student Individual or Group Project

- Students will choose a sub-topic of AI, write a paper, and make a class presentation

Appendix B: Artificial Intelligence Terms Glossary

Adversarial Examples: A security threat unique to machine learning. It is the intentional inclusion of an example to mistrain an AI. Adversarial Examples would cause AI to develop a vulnerability, for example an Autonomous Vehicle could be trained to not recognize a stop sign.

Algorithms: The "gears" that make Artificial Intelligence work. A set of instructions or mathematical functions that deliver a command or output from interpretation and analysis of inputted data. Algorithms can perform calculations, data processing and analysis, classification, and reasoning, to name a few

Artificial Intelligence: A blanket term which describes technologies that allow computers and machines to perform tasks that require human intelligence or capabilities like visual perception, decision making, reasoning, and translation. Examples of these technologies include machine learning and neural networks. AI can scour massive amounts of data to detect patterns in a short amount of time.

Augmented Intelligence: The current state of Artificial Intelligence. Human ability has been amplified by the use of these technologies. Pattern recognition and information processing are outsourced to computers and made much quicker than could ever be done by humans. Humans can review recommendations or analysis by computers to aid in decisions or other work. Instead of replacing humans, these technologies compliment their abilities by making them faster. However, the machines are still not "intelligent" in themselves in needing to be directed in their learning and their generally narrow applications.

Blockchain: An emerging system for digital storage that relies on decentralization and encryption to make stored data more resistant to attack and keep records safe. Some have advocated developing blockchain alongside AI to keep sensitive data more secure, help researchers follow algorithmic thought processes (which are currently very difficult or impossible to ascertain), and because AI will be able to manage blockchains better than humans.

Chatbot: AI technologies that are able to communicate with humans, usually through computer text. Commonly, they are deployed as customer service reps to handle consumer complaints and either offers a simple answer, assist in navigation of a website, or directing the customer to a proper representative.

Classification: Process of categorizing a group of objects while only using some data features about objects. It is a common application of AI. For example, identifying a picture as containing a dog or not, identifying applications that should pass or be rejected, and flagging outliers in datasets like tax records to detect fraud.

Data Fusion: merging of multiple heterogeneous datasets into one homogenous representation to be better processed for data mining and management

Data mining: analysis of large datasets to look for relationships between variables. Often this is done in an automated way where a computer does a rapid succession of tests to identify relationships. Humans must judge the substantive value and validity of those relationships.

Data Silos: The phenomena where different government agencies have troves of data that is isolated within that agency, disconnected from the data of other agencies. If the data was more integrated and connected, more insights, and more valuable algorithms could be developed.

Deep Learning: any artificial neural network that learn a long chain of causal links; more complicated neural networks with more nodes ("neurons") and connections between them. AI with more advanced neural networks can go beyond learning specific tasks and also identify the features of the task. This creates a better understanding of what is being done by the algorithm, and may allow it to generalize its task or more effectively improve its performance. It is described by some as the AI learning to ask "why". For example, instead of simply learning to identify pictures with dogs, deep learning will identify fur, paws, and eyes as separate components of a dog.

Expert Systems: An application of AI. The AI is programmed to make decisions that typically require human level expertise in things like making medical diagnosis or driving a car. These technologies are largely behind automation.

Features: Any identifying characteristic AI uses to analyze, compare, or make decisions. In visa applications this may be a place of origin or criminal history; word choice for natural language processing; or eye color in facial identification.

General Intelligence: Human level intelligence which is able to do things like make connections between tasks, combine skills, become cognizant, and other more complex abilities. The dream of AI would be to develop a program that could reach this level of cognition. Currently, AI systems are unable to operate outside of their narrowly defined tasks and are not considered "Generally Intelligent".

Inputs: AI requires a way to perceive its spatial or analytic environment, depending on the functions; this is the data it uses. Visual, audible, numerical, or qualitative data can be the input and it can be structured or unstructured. But not all AI can take all data, it depends on what it has been programmed to analyze.

Machine Learning: The process a machine goes through to teach itself a task and improve its completion of the task. Data is input, analyzed, and an output is returned. The machine can then compare its performance with a training baseline or its standalone performance (supervised or unsupervised learning) and improve its performance without human intervention. But, there will likely human oversight to evaluate its performance. In other words, the machine creates a (statistical) model of what it is doing or investigating and refines that model as more information is given to it. For example, the model may learn that fur is not a strong identifying characteristic of dogs after it wrongly uses fur to classify a cat as a dog.

Natural Language Processing: Computer translating and understanding of human language. For example, your iPhone personal assistant hearing your voice, understanding what you requested, and returning results or performing voice to text functions. Other examples include translators and chatbots.

Nearest Neighbor: Machine Learning where the algorithm compares its performance to the most similar observation, its 'nearest neighbor', or nearest K^{th} neighbors depending on the researchers' desires. Typically, this works in the realm of classification, like Amazon's recommendations or ad services. To recommend a new movie, Netflix will identify the 'nearest neighbors' to the movie you just rated 5 stars.

Neural Networks: A specific technology of AI that tries to simulate communicating brain neurons and allows it to learn. Input is split into separate simulated neurons which given weights for their importance in analysis. The neurons perform a process and "score" their data. In other terms this means the input is analyzed, and the information is sent to other neurons for other analysis until it reaches the output, where a decision is made. The adjustment of weights given to each node is part of the learning process. Nodes must determine where to pass along the information they are given, which comprises how they are connected within the neural network. Output nodes may be classifiers, like accept or reject. In this case, the node with the higher score is the deciding one. The middle notes would identify if the information matches what it knows about passing or not passing observations and scores based on that.

Outputs: AI output can be broad. From real-time traffic data the output may be a red or green light on a traffic signal. In visa applications, the AI can return recommendations like accept or reject. The AI can return correlations or patterns in data to flag outliers or provide new information. It may also be a response to a customer question. The output that AI can perform is diverse.

Propagation: The process of working through a neural network. There is forward propagation, or moving from the inputs to the outputs. There is also backward propagation, where the machine works backwards and assesses its process and can adjust the weights or making other changes to refine itself, this is how it learns or trains itself.

Reinforcement Learning: Machine learning where the algorithm checks its performance against a "correct" example given by a human. It can identify if it was right or wrong and adjust based on that outcome. It can also identify numerous solutions for humans to review. Researchers can set rewards or punishments for behavior can be set to direct the development of the algorithm, and the algorithm with then seek to optimize itself by minimizing punishment and maximizing reward. One must be careful in setting these constraints as they will be taken very literally and precisely by the computer.

Robotics: A field of AI which incorporates hardware (robots) into the 'thinking' software.

Robotic Process Automation: transferring human tasks over to machines. This often involves highly repetitive tasks which robots can do more efficiently.

Social Intelligence or Affective Computing: Systems that recognize, interpret, or process human affects like emotions, vocal tone, facial expression, etc.

Structured Data: information with a high degree of organization, such that inclusion in a relational database is seamless and readily searchable by simple, straightforward search engine algorithms or other search operations

Supervised Learning: Machine Learning where example inputs to outputs are given for reference to the algorithm. The algorithm learns how to replicate those connections or reach those conclusions. Humans have intervened to direct the algorithm.

Training: In Machine Learning, it is the process of improving the machines performance through running it with more data; usually the algorithm must be given a training dataset that contains the "right" answer for it to learn from initially, but later data sets can omit the answer to evaluate its performance. More and better data will help construct a better model.

Turing Test: A test developed by British AI pioneer Alan Turing in the 1950's to determine if a technology is intelligent. Simply it says that if a human interacting with a machine is unable to identify it as a machine and not a human, then that machine is intelligent. This test is often applied to chatbots.

Unstructured Data: Data that is poorly organized so to be hard for us to understand it or search within it. Data may be unlabeled or uncategorized, or poorly so. A messy email inbox may be an example of unstructured data because while data are labeled by sender and subject, they may be disorganized within the inbox.

Unsupervised Learning: Machine learning where no example inputs to outputs or connections are identified for the algorithm. Instead, it independently seeks to identify connections or patterns. An application would be data mining in a large dataset.

Weighting: Giving different importance levels to different bits of information. More important information is weighted more heavily. In neural networks, the nodes that perform analysis of different information are weighted based on how much they contribute to the final decisions.

Appendix C: Abstracts for Articles Relating to Ethics of AI

Prepared by: Caroline Ball (MPA student, George Mason University) and Anders Shropshire (Academy Intern, University of Wisconsin-Madison)

Prepared for: Academy AI Working Group - Subgroup on Ethics and Bias

Association for Computing Machinery. 2017. "**Statement on Algorithmic Transparency and Accountability**" (ACM U.S. Public Policy Council, approved January 12, 2017 and ACM Europe Policy Committee, approved May 25, 2017). https://www.acm.org/binaries/content/assets/public-policy/2017_joint_statement_algorithms.pdf

The ACM approved seven principles for algorithmic accountability and transparency, including: awareness; access and redress; accountability; explanation; data provenance; auditability; validation and testing. This set of principles, consistent with the ACM Code of Ethics, is intended to support the benefits of algorithmic decision-making while addressing concerns about discrimination, opaqueness, and error.

Barocas, Solon & Selbst, Andrew. 2016. "**Big Data's Disparate Impact**," *California Law Review* vol. 104 (issue 3): 671-732. http://www.californialawreview.org/wp-content/uploads/2016/06/2Barocas-Selbst.pdf

This California Law Review examines concerns of inherent discrimination within data mining through the lens of Title VII's prohibition of discrimination in employment. Despite advocates' beliefs, algorithms are not a perfect tool because they are as biased as the data that are fed in and often reflect historic patterns of injustice, unintentional though they may be. There are technical, legal, and political difficulties in remedying the situation and the best efforts of those involved may not be enough – a reexamination of the historical and modern meanings of "discrimination" and "fairness" will be required.

Brauneis, Robert & Goodman, Ellen. 2017. "**Algorithmic Transparency for the Smart City**," *Yale Journal of Law and Technology* vol. 20: 103-176. https://ssrn.com/abstract=3012499

This article examines the implications of Smart Cities in both an ethical and legal context, with attention to transparency in the deployment and development of predictive algorithms. The main focus of the article is on whether or not the public has the ability to learn about and independently assess predictive algorithms being deployed around the country. They seek information on government algorithms through Freedom of Information requests. They find that for the general public, getting access to documents relating to the development of algorithms, as well as the algorithms themselves, is difficult or impossible. Much of the article goes through case studies of the authors attempting to get access to this information through FOIA requests. The authors suggest there are ethical issues from first, a potential misalignment of values in the local government's goals and the encoded prerogative of the algorithm. Secondly, if the public administrators do not fully understand the workings of the algorithm, there is no way for them to be held accountable to the public.

Campolo, Alex; Sanfilippo, Madelyn; Whittaker, Meredith; Crawford, Kate. 2017. "**2017 Report**," *AI NOW*. https://ainowinstitute.org/AI_Now_2017_Report.pdf

This year-end report compiled by experts and researchers of Artificial Intelligence warns that even though the concept of AI has been around for decades, real-world applications of it are now careening into prominence and require careful consideration. The report includes recommendations for companies and agencies using AI, all in the vein of rigorous preparation before use and post-implementation monitoring. The four major problem areas identified for the progress of AI are labor and automation, bias and inclusion, rights and liberties, and ethics and governance. The report concludes that a better understanding of AI and developing ethical codes for its use are paramount to its success.

Center for Internet and Human Rights. 2015. "**The Ethics of Algorithms: From Radical Content to Self-Driving Cars**," *Centre for Internet and Human Rights.* https://cihr.eu/wp-content/uploads/2015/07/Ethics_of_Algorithms_Final_Paper.pdf

In this 2015 report from the Centre for Internet and Human Rights, the issue of ethical considerations for algorithms is tackled. Discrimination and the balance of power or lack thereof can all be magnified using algorithms and often are done so unknowingly. There are three specific attributes of algorithms that require ethical scrutiny – complexity and opacity, gatekeeping functions, and subjective decision-making. As usual, the authors provide suggestions for legislative response and regulation.

Citron, Danielle Keats. 2008. "**Technological Due Process**," *Washington University Law Review* vol. 85 (issue 6): 1249-1313. https://openscholarship.wustl.edu/cgi/viewcontent.cgi?article=1166&context=law_lawreview

This article focuses on the modern data mining process's skill at skirting around due process and not allowing individuals and the legal system to dictate the rules it follows. It draws a contrast between the twentieth and twenty-first century automation systems and how regulation has not had a chance to catch up and put adequate rules in place. The author calls for an updated framework through which the process of data mining can be held accountable in the same respect than any human process is.

Citron, Danielle Keats & Pasquale, Frank. 2014. "**The Scored Society: Due Process for Automated Predictions**," *Washington University Law Review* vol. 89: 1-33. https://digital.law.washington.edu/dspace-law/bitstream/handle/1773.1/1318/89wlr0001.pdf?sequence=1

In this article, credit scores are used as a case study to demonstrate the potential negative consequences of using predictive scores derived from algorithms to guide actionable decisions. The authors contest the claim that these predictions are objective by reminding that human judgement and values helped create the algorithms. These systems have the ability to cause a great deal of harm, like locking individuals out of access to credit, thus becoming a self-fulfilling prophecy. They raise concerns of algorithms having a disparate impact on protected classes. In order to ensure consumer and citizen protection from a new era of

discrimination, the authors recommend a system of due process for citizens that provide procedural safe guards at each step of the scoring process; collection, calculation, dissemination and use.

Crawford, Kate & Schultz, Jason. 2014. "**Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms**," *Boston College Law Review* vol. 55 (issue 1): 93-128. https://lawdigitalcommons.bc.edu/cgi/viewcontent.cgi?article=3351&context=bclr

There is an inherent dilemma in the modern practice of data mining – even if the data gathered is to the benefit of the consumer, how can that consumer's privacy still be protected? The authors state that data mining has rapidly increased the amount of "personally identifiable information" of an individual and that these can render inaccurate profiles that nonetheless negatively impact the individual. The article discusses various privacy concerns and the three legal acts from the 1970s-80s that usually address these concerns, but then claims that these acts are inadequate to handle the current privacy problems with big data. The regulation that the authors recommend would be most effective is that of "procedural data due process", the same due process exemplified in the American judicial system.

Desai, Deven & Kroll, Joshua A. 2018. "**Trust But Verify: A Guide to Algorithms and the Law**," *Harvard Journal of Law and Technology* vol. 31 (issue 1): 1-64. https://jolt.law.harvard.edu/assets/articlePDFs/v31/31HarvJLTech1.pdf

This comprehensive guide to algorithms by the Harvard Journal of Law and Technology describes the basic premise and issues with the use of algorithms, both in the private and public sectors and even the decision of where and when to use them. There is a focus on transparency required for those implementing practices with algorithms so that they are not repeatedly generating flawed outcomes. Computer science offers practical solutions for the problems that arise with algorithms in the form of white-box and black-box testing.

Edwards, Lilian & Veale, Michael. 2017. "**Slave to the algorithm? Why a right to explanation is probably not the remedy you are looking for**," *Duke Law and Technology Review* vol. 16 (issue 1): 1-65. *https://scholarship.law.duke.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1315&context=dltr*

This paper makes the case that a legally guaranteed "Right to Explanations" is not sufficient to resolve the problems and fears many have about automated decision making, and instead create an illusion of transparency. The author suggests that many grievances about the role of AI in decision making would not be resolved by this right because learning about the decision-making criteria would not affect the perverse outcome. In general, other protections derived in the GDPR provide a more promising foundation for protecting consumers from harm. Additionally, the author notes the technical barriers standing in the way and suggests that such a right will only hamper efforts to further develop and improve these technologies.

Executive Office of the President. 2014. "**Big Data: Seizing Opportunities, Preserving Values**." https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf

In May of 2014, members in the Executive Office of the President submitted the results of a 90-day study that the President had requested they perform on big data and its effects. After a description of big data, how it affects individuals, and the Obama administration's general approach, the report launches into an examination of big data use in the public sector, private sector, and policy frameworks that would affect both. The conclusion focuses on the need to retain privacy policies as well as several policy enforcement recommendations.

Federal Trade Commission. 2016. "**Big Data: A Tool for Inclusion or Exclusion**," *Federal Trade Commission.* https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf

In 2016, the FTC held a workshop under this same title and subsequently compiled the information from the workshop and other research into a report on the expanding role and implications of big data. After a brief description of the life cycle of big data, the report focuses specifically on its use and the issues that can arise when biases are introduced. The FTC outlines laws that would affect companies using big data and the current policy considerations that may affect them as well. Also noted is the duality of the issue of gathering and using big data in the first place – gathering information to benefit consumers while also somehow protecting consumer privacy.

Hardt, Moritz; Price, Eric; Srebro, Nathan. 2016. "**Equality of Opportunity in Supervised Learning**," *30th Conference on Neural Information Processing Systems.* **https://arxiv.org/pdf/1610.02413.pdf**

This highly technical paper develops a framework to assess and reduce bias in predictive algorithms. It creates precise definitions of Equality of Opportunity and Equalized odds. The authors create a framework that is simple and does not depend on analyzing each characteristic of a case, but rather the joint distribution of outcomes and protected status through oblivious measures. In addition to building their own framework, the authors briefly explain the strengths and weaknesses of past work on eliminating bias in AI. To illustrate the different trade-offs and outcomes, the authors use loan approval and credit scoring to show how different criteria dramatically affect requirements.

Hiller, Janine & Blanke, Jordan. 2017. "**Smart Cities, Big Data, and the Resilience of Privacy**," *Hastings Law Journal* vol. 68 (issue 2): 309-356. *http://www.hastingslawjournal.org/wp-content/uploads/Hiller_Blanke-68.2.pdf*

This article looks to provide a theory for how the privacy system will adapt to the growing penetration of data into our lives and includes interesting histories about the development of privacy and methods and purpose of data collection in Smart Cities. The authors examine how resilient existing frameworks and regulatory approaches are for protecting privacy, concluding that they may suffer from a lack of citizen

input and adaptability. In the discussions, they highlight how Smart City data collection regimes conflict with past Supreme Court cases regarding protection from invasive data collection but note that SmartCities have not yet been challenged by case law.

Kim, Pauline T. 2017. "**Auditing Algorithms for Discrimination**," *University of Pennsylvania Law Review* vol. 166 (issue 1): 189-203. https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=1212&context=penn_law_review_online

This essay serves as a response to Kroll et al. and their claims that transparency is not a solution to algorithm accountability. Kim argues the opposite, saying that transparency in the form of third-party auditing offers the clearest assessment of the effectiveness and potential discriminatory nature of any algorithm.

Kim, Pauline T. 2017. "**Data-Driven Discrimination at Work**," *William and Mary Law Review* vol. 58 (issue 3): 857-936. https://core.ac.uk/download/pdf/83126733.pdf

The focus of this article lies within the employment sphere, specifically that of using algorithms in the hiring process and the discrimination that can arise. The author echoes the concerns of those wary about the use of algorithms – that if algorithms are built with inaccurate data, the biases inherent in that data will pass along to and be magnified by the algorithm. The author identifies the biases that result from inaccurate hiring algorithms as classification bias and states that it this is prohibited by Title VII and requires legal ramifications.

Kroll, Joshua A. et al. 2017. "**Accountable Algorithms**," *University of Pennsylvania Law Review* vol. 165 (issue 1): 633-705. https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=9570&context=penn_law_review

The authors of this article recognize that there is much to be desired for the accountability of data mining and use and that transparency is the most popular solution. They argue, however, that this is often neither necessary nor sufficient to solve the problem and that there are other technological routes that would still preserve the privacy of those whose information is being mined. The article describes in-depth these various routes and how they both comply with legal requirements and privacy concerns.

Lepri, Bruno et al. 2016. "**The Tyranny of Data? The Bright and Dark Sides of Data-Driven Decision-Making for Social Good**," *Transparent Mining for Big and Small Data* vol. 32: 3-24. *https://arxiv.org/pdf/1612.00323.pdf*

This article focuses on algorithms designed to optimize resource allocation in efforts to better the lives of individuals, commonly referred to as social good algorithms. *The authors overview the applications and benefits of these algorithms before raising issues with privacy violations, informational asymmetry, lack of transparency, discrimination and social exclusion which can easily result from these tools. The authors*

then introduce "human-centric" requirements for positive disruption. They make recommendations for user-centric data ownership, tools and participatory infrastructure to encourage transparency, and developing living labs to experiment with and create data driven policies.

Mehr, Hila. 2017. "**Artificial Intelligence for Citizen Services and Government**," *Harvard Ash Center for Democratic Governance and Innovation.* https://ash.harvard.edu/files/ash/files/artificial_intelligence_for_citizen_services.pdf

This paper asks how we can use AI to improve government for citizens and rebuild trust. The focus of the article is on applications that directly relate to citizen services like answering questions, filling out and searching documents, routing requests, translation, and drafting documents. The author makes the following recommendations for agencies looking to implement AI - make AI a part of a citizen centric program, get citizen input, build upon existing resources, be data-prepared and tread carefully with privacy, mitigate ethical risks and avoid AI decision making, and augment employees but do not replace them.

Mittelstadt, Brent Daniel et al. 2016. "**The Ethics of Algorithms: Mapping the Debate**," *Big Data and Society* July-December: 1-21. http://journals.sagepub.com/doi/pdf/10.1177/2053951716679679

This article provides overviews of the most concerning ethical issues related to decision-making algorithms becoming more ubiquitous in society. The authors note how values are inescapably coded into algorithms and explain how difficult it is to identify bias in algorithms and assess its extent. The authors contribute to the debate by providing a structure for discussing future issues that arise based on 6 categories of ethical concerns; inconclusive evidence, inscrutable evidence, misguided evidence, unfair outcomes, transformative effects, and traceability.

Newell, Sue & Marabelli, Marco. 2015. "**Strategic Opportunities (and challenges) of Algorithmic Decision-Making: A Call for Action on the Long-term Societal Effects of 'Datification'**," *Journal of Strategic Information Systems* vol. 24 (issue 1): 3-14. **https://www.researchgate.net/publication/273478334**

This article examines ethical trade-offs inherent in the deployment of artificial intelligence and the growth of big data. They identify three non-exhaustive tradeoffs as privacy vs security, freedom vs control, and independence vs dependence. They ask about the effects of internet bubbles on people's world views, decision making and learning when data is not viewed by people, and people's respect for a computer's decision. The authors identify two schools of moral thought for considering these questions.

Pitoura, Evaggelia et al. 2017. "**On Measuring Bias in Online Information**," *Association for Computing Machinery's Special Interest Group on Management of Data* vol. 46 (issue 4): 16-21. *https://arxiv.org/pdf/1704.05730.pdf*

The authors of this brief article aim to present a systematic approach to measuring bias in algorithms in the context of online information. This includes search results in places like Google and Amazon, and content presented to consumers on social media and elsewhere. The article provides background on the problems associated with information bias and break it down into core components and influences. The authors differentiate between user bias and content bias and explain differences between group level fairness and individual level fairness.

President's Council of Advisors on Science and Technology, 2014. "**Big Data and Privacy: A Technological Perspective**," *Executive Office of the President.* https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf

This White House report offers a thorough overview of changes expected to be brought about by the Big Data revolution, including many related to ethics of data collection, storage, and use in respect to privacy. While much of the report focuses on the technical aspects and technological development of big data, it also explains legal precedents and social expectations of privacy and how they are strained by both past and expected advancements. The report brings to light the complicated trade-offs between data collection and privacy, with the authors noting they believe the benefits of big data will outweigh the negatives.

Rainie, Lee & Anderson, Jana. 2017. "**Code-Dependent: Pros and Cons of the Algorithm Age**," *PEW Research Center.* http://www.pewinternet.org/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age

Pew Research surveyed experts in the field of AI to gauge the overall effect of algorithms in society, receiving many answers that identify ethical questions about the future we are heading towards. This article provides a survey of the landscape of ethical concerns in AI with direct quotations from leading experts. Those experts expressed concern in giving robots decision-making power because of many issues including biases from coders' social biases and historical bias in data, transparency issues in the development and operation of AI, granting institutions unseen levels of power for manipulation provided by analytic insights, and classification becoming a self-fulfilling prophecy in some ways and promoting echo chambers in others.

Regan, Priscilla M. 2017. "**Big Data and Privacy**," in *Analytics, Policy and Governance*. Ed by Jennifer Bachner, Kathryn Wagner Hill, and Benjamin Ginsberg. New Haven: Yale University Press.

This chapter discusses the conceptualization of the privacy problems posed by "big data" based on how data practices and analytics are currently, or projected to be, used in a number of contexts. Privacy problems include controlling collection and use of information about oneself, autonomy over decision making, anonymity, choice in group associations, and "practical obscurity"—as well as related values of due process, equal protection, data security, and accountability. The chapter then identifies and evaluates a number of policy approaches for addressing these problems, demonstrating that the power of big data renders traditional information privacy policy responses obsolete and ineffective and argues for regulation and oversight on entities collecting and using big data.

Regan, Priscilla M and Jolene Jesse. 2018. "**Ethical challenges of edtech, big data and personalized learning: twenty-first century student sorting and tracking**," *Ethics and Information Technology* (published online December 3). https://doi.org/10.1007/s10676-018-9492-2

This article analyzes the range of these ethical concerns arguing that characterizing them under the general rubric of 'privacy' oversimplifies the concerns and makes it too easy for advocates to dismiss or minimize them. Six distinct ethical concerns are identified: information privacy; anonymity; surveillance; autonomy; non-discrimination; and ownership of information. Particular attention is paid to whether personalized learning programs raise concerns similar to those raised about educational tracking in the 1950s.

Salidek, Adam et al. 2016. "**Deploying nEmesis: Preventing Foodborne Illness by Data Mining Social Media**," *Association for the Advancement of Artificial Intelligence* Twenty-Eighth AAAI Conference: 3982-3989. https://www.aaai.org/ocs/index.php/IAAI/IAAI16/paper/view/11823

This article outlines a deployment of AI that screens social media posts in Las Vegas to identify restaurants that are causing illness and may be violating health code laws. They use this information to target health inspections, improve efficiency, and identify more problem establishments. This novel application of data collection and analysis raises questions about the degree of state surveillance that should be allowed or desired. Crime, public opinion, and signs of welfare more may be effectively and passively monitored by governments with data collection and AI analysis, but this article centers on the unsettling thought of the government monitoring citizens' social media.

Sweeney, Latanya. 2013. "**Discrimination in Online Ad Delivery**," *ACM Queue* vol. 11 (issue 3): 1-19. https://queue.acm.org/detail.cfm?id=2460278

This article examines the issue of racial discrimination stemming from online advertisements, specifically which names the advertisement pairs with the term "arrest". Through extensive searches on Google, Google Images, and seeing what ads Google AdSense displays, Sweeney's research finds that there is a correlation in the delivery of ads suggestive of an arrest record based on searches of racially associated names and promises future research and development on the topic.

Tene, Omer & Polonetsy, Jules. 2013. "**Big Data for All: Privacy and User Control in the Age of Analytics**," *Northwestern Journal of Technology and Intellectual Property* vol. 11 (issue 5): 239-273. https://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?article=1191&context=njtip

This article serves as an overview of the current state of big data and privacy concerns. The practicalities of data mining and its benefits to an individual consumer are described in depth, as are the concerns to privacy that this access to personal information poses. The legal framework for legislating privacy in the wake of the explosion of data mining is examined in the same fashion, with the problems and possible solutions discussed equally.

Appendix D: Bibliography

Agrawa, Ajay et. al. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Review Press, 17 Apr. 2018.

*AI for Good Global Summit Repo*rt. United Nations, Xprize, ITU, 2017, www.itu.int/en/ITU-T/AI/Documents/Report/AI_for_Good_Global_Summit_Report_2017.pdf.

Aiyaswami, Mohan. "Smarter Australian Defence Force the Information Advantage – CEBIT." Technology, May 2016. https://www.slideshare.net/CeBITAustralia/cebitaustralia2016mohanaiyaswamiegovernmenttransformationdefence.

Angwin, Julia, et al. "Machine Bias." *ProPublica,* 23 May 2016, www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Barocas, Solan, and Andrew D Selbst. "Big Data's Disparate Impact." *California Law Review*, vol. 104, no. 3, 11 Aug. 2016, pp. 671–732., doi:10.18411/d-2016-154.

Brauneis, Robert and Goodman, Ellen P., Algorithmic Transparency for the Smart City (August 2, 2017). 20 Yale J. of Law & Tech. 103 (2018); GWU Law School Public Law Research Paper; GWU Legal Studies Research Paper. Available at SSRN: https://ssrn.com/abstract=3012499 or http://dx.doi.org/10.2139/ssrn.3012499.

Bright Planet. "Structured vs. Unstructured data." Brightplanet.com. https://brightplanet.com/2012/06/structured-vs-unstructured-data/ (accessed 3/22/19)

Broussard, Meredith. *Artificial Unintelligence: How Computers Misunderstand the World.* The MIT Press, 27 Apr 2018.

Brundage, Miles, et al. *The Malicious Use of AI: Forecasting, Prevention, and Mitigation*. University of Oxford, 2018, *The Malicious Use of AI: Forecasting, Prevention, and Mitigation*, arxiv.org/ftp/arxiv/papers/1802/1802.07228.pdf.

Bucher, Taina. *If…Then: Algorithmic Power and Politics.* Oxford University Press, 26 June 2018

Burrell, Jenna. *How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms*. Big Data and Society, 2016, journals.sagepub.com/doi/pdf/10.1177/2053951715622512.

Carton, Samuel. et al. "Identifying Police Officers at Risk of Adverse Events." Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16 (New York, NY, USA: ACM, 2016), 67–76, doi:10.1145/2939672.2939698.

Chun, Andy, and Hon Wai. "AI Success Stories." *Dr. Andy Chun - Artificial Intelligence Success Stories, City University of Hong Kong*, www.cs.cityu.edu.hk/~hwchun/AIProjects/stories/.

Chun, Andy. *Using AI for e-Government Automatic Assessment of Immigration Application Forms*. 2007, www.cs.cityu.edu.hk/~hwchun/research/PDF/iaai_2007.pdf.

Chun, H.W., and Ted Suen, "Engineering Works Scheduling for Hong Kong's Rail Network," In Proceedings of the 26th Conference on Innovative Applications of Artificial Intelligence Conference, Quebec City, Quebec, Canada, 27-31 July, 2014.

Citron, Danielle Keats and Frank A. Pasquale, "The Scored Society: Due Process for Automated Predictions" (2014). Washington Law Review, Vol. 89, 2014, p. 1-; U of Maryland Legal Studies Research Paper No. 2014-8. Available at SSRN: https://ssrn.com/abstract=2376209

Citron, Danielle Keats. "Technological Due Process." *Washington University Law Review*, vol. 85, no. 6, 2008, pp. 1249–1313., doi:10.18411/d-2016-154.

Contizer, Vincent, et al. "Moral Decision Making Frameworks for Artificial Intelligence." *Duke University,* Association for the Advancement of Artificial Intelligence, 2017, users.cs.duke.edu/~conitzer/moralAAAI17.pdf.

Crawford, Kate, and Jason Schultz. "Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms." *Digital Commons @ Boston College Law School*, 29 Jan. 2014, lawdigitalcommons.bc.edu/bclr/vol55/iss1/4/.

Desouza, Keith C. "Delivering Artificial Intelligence in Government: Challenges and Opportunities," IBM Center for the Business of Government, 2018, businessofgovernment.org/sites/default/files/Delivering%20Artificial%20Intelligence%20in%20Government.pdf.

Desouza, Kevin C., and Kiran Kabtta Somvanshi. "How Blockchain Could Improve Election Transparency." *Brookings*, Brookings Institution, 29 May 2018, www.brookings.edu/blog/techtank/2018/05/30/how-blockchain-could-improve-election-transparency .

Dzone. "28 Artificial Intelligence Terms You Need to Know." Dzone.com. https://dzone.com/articles/ai-glossary (accessed 3/22/19).

Etzioni, Amitai and Oren Etzioni. "Should Artificial Intelligence Be Regulated?" *Issues in Science and Technology*, 2017, pp. 32–36, ai2-website.s3.amazonaws.com/publications/Etzioni_Regulate_AI.pdf.

Etzioni, Oren. "How to Regulate Artificial Intelligence." *The New York Times*, 2 Sept. 2017, www.nytimes.com/2017/09/01/opinion/artificial-intelligence-regulations-rules.html.

Eubanks, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* St. Martin Press, 23 Jan 2018.

Frischmann, Brett and Evan Selinger. *Re-Engineering Humanity.* Cambridge University Press, 19 Apr 2018

Goodman, Bryce, and Seth Flaxman. "European Union Regulations on Algorithmic Decision-Making and a 'Right to Explanation.'" Cornell University Library, 31 Aug. 2016, arxiv.org/abs/1606.08813.

Governing Machine Learning: Exploring the Intersection between Machine Learning, Law, and Regulation. Yale Law School, Immuta, 2017, law.yale.edu/system/files/area/center/isp/documents/governing_machine_learning_-_final.pdf.

Halpern, Joe. "Moral Responsibility, Blameworthiness, and Intention: In Search of Formal Definitions." CS 4732. Ethical and Social Issues in AI, 1 May 2017, Ithaca, NY, Cornell University, www.cs.cornell.edu/courses/cs4732/2017sp/.

Hardt, Mortiz, et al. "Equality of Opportunity in Supervised Learning." Cornell University Library, 7 Oct. 2016, arxiv.org/abs/1610.02413.

Harkinson, Josh. "Could This Algorithm Pick Which Homeless People Need Housing Most Urgently?" *Mother Jones*, 23 June 2017, www.motherjones.com/politics/2016/06/homelessness-data-silicon-valley-prediction-santa-clara/.

Heaton, Brian. "New York City Fights Fire with Data." *Government Technology: State & Local Government News Articles, Government Technology*, 15 May 2015, www.govtech.com/public-safety/New-York-City-Fights-Fire-with-Data.html.

Hiller, Janine S, and Jordan M Blanke. "Smart Cities, Big Data, and the Resilience of Privacy." *Hastings Law Review*, vol. 68, no. 2, Feb. 2017, pp. 309–356.

Husain, Amir. *The Sentient Machine: The Coming Age of Artificial Intelligence.* Scribner, 21 Nov. 2017.

Kamensky, John. "Civic Engagement in the Future." *IBM Center for the Business of Government*, 22 May 2018, businessofgovernment.org/blog/civic-engagement-future.

Kamensky, John. "Government Is Data Rich, But Information Poor." Government Executive, 12 June 2018, www.govexec.com/excellence/management-matters/2018/06/government-data-rich-information-poor/148914/.

Kamensky, John. "The Future of Work in Government." *IBM Center for the Business of Government*, 27 Feb. 2018, businessofgovernment.org/node/2641.

Kamensky, John. "The Role of Artificial Intelligence in the Future of Government." *IBM Center for the Business of Government*, 19 Apr. 2018, businessofgovernment.org/node/2686.

Kleinberg, Jon, et al. "A Guide to Solving Social Problems with Machine Learning." Harvard Business Review, 8 Dec. 2016, hbr.org/2016/12/a-guide-to-solving-social-problems-with-machine-learning.

Kleinberg, Jon, et al. *Human Decisions and Machine Prediction*. NBER, 2017, www.nber.org/papers/w23180.

Kleinberg, Jon. "Inherent Trade-Offs in Algorithmic Fairness." CS 4732. 6 March 2017, Ithaca, NY, Cornell University, www.cs.cornell.edu/courses/cs4732/2017sp/.

Knepper, Ross. "Autonomy, Embodiment, and Anthropomorphism: The Ethics of Robotics." CS 4732. Ethical and Social Issues in AI, 24 Apr. 2017, Ithaca, NY, Cornell University, www.cs.cornell.edu/courses/cs4732/2017sp/.

Leben, Derek. *Ethics for Robots.* Routledge (27 June 2018).

Lepri, Bruno, et al. "The Tyranny of Data? The Bright and Dark Sides of Data-Driven Decision-Making for Social Good." *Cornell University Library*, 2 Dec. 2016, arxiv.org/abs/1612.00323

Levy, Karen. "Working with and Against AI: Lessons from Low-Wage Labor." CS 4732. Ethical and Social Issues in AI, 17 Apr. 2017, Ithaca, NY, Cornell University, www.cs.cornell.edu/courses/cs4732/2017sp/.

LohmeyerJul, Suzette. "Faster and More Accurate Flood Prediction." *GCN,* 21 July 2016, gcn.com/articles/2016/07/21/stormsense.aspx.

Luca, Michael, et al. "Algorithms Need Managers, Too." Harvard Business Review, 2016, hbr.org/2016/01/algorithms-need-managers-too.

Madaio, Michael, et al. "Identifying and Prioritizing Fire Inspections: A Case Study of Predicting Fire Risk in Atlanta." 2015. https://www.cc.gatech.edu/~bdilkina/papers/madaio2015identifying.pdf.

Madaio, Michael, et al. *Firebird: Predicting Fire Risk and Prioritizing Fire Inspections in Atlanta*. 2016, Firebird: Predicting Fire Risk and Prioritizing Fire Inspections in Atlanta, firebird.gatech.edu/KDD16_Firebird.pdf.

Markoff, John. "As Artificial Intelligence Evolves, So Does Its Criminal Potential." The New York Times, The New York Times, 23 Oct. 2016, www.nytimes.com/2016/10/24/technology/artificial-intelligence-evolves-with-its-criminal-potential.html.

Marr, Bernard. *Artificial Intelligence and Blockchain: 3 Major Benefits of Combining These Two Mega-Trends*. Forbes, Mar 2, 2018. https://www.forbes.com/sites/bernardmarr/2018/03/02/artificial-intelligence-and-blockchain-3-major-benefits-of-combining-these-two-mega-trends/#4bf90e044b44 (accessed March 22, 2019).

Mayer-Schonberger, Viktor. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Mariner Books, 4 Mar. 2014.

Mehr, Hila. *Artificial Intelligence for Citizen Services and Government*. Harvard Ash Center, 2017, ash.harvard.edu/files/ash/files/artificial_intelligence_for_citizen_services.pdf.

Mittelstadt, Brent Daniel, et al. "The Ethics of Algorithms: Mapping the Debate." *Big Data & Society*, vol. 3, no. 2, 1 Dec. 2016, doi:10.1177/2053951716679679.

National Science and Technology Council Committee on Technology. Preparing for the Future of Artificial Intelligence. 2016, obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

Newell, Sue, and Marco Marabelli. "Strategic Opportunities (and Challenges) of Algorithmic Decision-Making: A Call for Action on the Long-Term Societal Effects of 'Datification.'" *Journal of Strategic Information Systems*, vol. 24, no. 1, Feb. 2015, doi:10.1016/j.jsis.2015.02.001.

Olteanu, Alexandra, et al. "Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries." *SSRN*, 20 Dec. 2016, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2886526.

Osoba, Osende A, and William Welser. *The Risks of Artificial Intelligence to Security and the Future of Work*. Rand Corporation, 2017, www.rand.org/pubs/perspectives/PE237.html.

Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information.* Harvard University Press, 5 Jan. 2015.

Peppet, Scott R." Regulating the Internet of Things: First Steps Toward Managing Discrimination, Privacy, Security & Consent," (March 1, 2014). *Texas Law Review*, Forthcoming. Available at SSRN: https://ssrn.com/abstract=2409074.

Pitoura, Evaggellia, et al. "On Measuring Bias in Online Information." *Cornell University Library*, 3 Oct. 2017, arxiv.org/abs/1704.05730.

President's Council of Advisors on Scienceand Technology. "Big Data and Privacy: A Technological Perspective." 2014, Big Data and Privacy: A Technological Perspective, obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_big_data_and_privacy_-_may_2014.pdf.

Rainie, Lee, and Janna Anderson. "Code-Dependent: Pros and Cons of the Algorithm Age." *Pew Research Center: Internet, Science & Tech*, 8 Feb. 2017, www.pewinternet.org/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/.

Rothe, Rasmus. "Applying Deep Learning to Real-World Problems – Merantix." *Medium*, Augmenting Humanity, 23 May 2017, medium.com/merantix/applying-deep-learning-to-real-world-problems-ba2d86ac5837.

Ruddick, Graham. "Facebook Forces Admiral to Pull Plan to Price Car Insurance Based on Posts." *The Guardian,* 2 Nov. 2016, www.theguardian.com/money/2016/nov/02/facebook-admiral-car-insurance-privacy-data.

Sadilek, Adam, et al. "Deploying NEmesis: Preventing Foodborne Illness by Data Mining Social Media." *Association for the Advancement of Artificial Intelligence*, vol. 38, no. 1, 2017, pp. 3982–3989., doi:https://doi.org/10.1609/aimag.v38i1.2711

Schwartz, Jeff, et al. "The Future of Work: The Augmented Workforce." *Deloitte United States*, 28 Feb. 2017, www2.deloitte.com/insights/us/en/focus/human-capital-trends/2017/future-workforce-changing-nature-of-work.html.

Scott, Tony, and Anne E Rung. "Federal Source Code Policy: Achieving Efficiency, Transparency, and Innovation through Reusable and Open Source Software." Received by Heads of Departments and Agencies, Executive Office of the President, 8 Aug. 2016, Washington, D.C.

See, Abigail. "Nearest Neighbors." Tutorial, SAILORS from Stanford AI4ALL, July 12, 2016, https://cs.stanford.edu/people/abisee/nn.pdf.

Selman, Bart. "The Future of AI: Benefits and Risks." CS 4732. 27 February 2017, Ithaca, NY, Cornell University, www.cs.cornell.edu/courses/cs4732/2017sp/.

Shark, Alan. "The Future Is Now: Artificial Intelligence and Robots and What It Means for Local Government." *Public Technology Institute*, 2018, www.pti.org/cals/previous/2018events/041818.asp.

Shubhendu, Shukla and Jaiswal Vijay. "Applicability of Artificial Intelligence in Different Fields of Life." (2013). https://www.semanticscholar.org/paper/Applicability-of-Artificial-Intelligence-in-Fields-Shubhendu-Vijay/c06ab259ede1ebefbd0ba01f0f4603c4b2bb19ea.

Smart Urban Signal Networks: Initial Application of the SURTRAC Adaptive Traffic Signal Control System. ICAPS (Citeseer, 2013). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.433.5935&rep=rep1&type=pdf.

Smith, Aaron, and Monica Anderson. "Automation in Everyday Life." *Pew Research Center: Internet, Science & Tech*, 4 Oct. 2017, www.pewinternet.org/2017/10/04/automation-in-everyday-life/.

Smith, Stephen F. *Smart Urban Signal Networks: Initial Application of the SURTRAC Adaptive Traffic Signal Control System*. Carnegie Mellon University Robotics Institute, June 2013, www.ri.cmu.edu/pub_files/2013/6/icaps2013.pdf.

Spielkamp, Matthias. "We Need to Shine More Light on Algorithms so They Can Help Reduce Bias, Not Perpetuate It." *MIT Technology Review*, 16 June 2017, www.technologyreview.com/s/607955/inspecting-algorithms-for-bias/.

StormSense Project. https://vims-wm.maps.arcgis.com/apps/MapJournal/index.html?appid=62c80853313743f3acf5a83ab420d015.

Sunstein, Cass. *The Ethics of Nudging*. Yale Journal on Regulation, 2015, https://digitalcommons.law.yale.edu/yjreg/vol32/iss2/6

Technology Assessment: Artificial Intelligence Emerging Opportunities, Challenges, and Implications. United States Government Accountability Office, 2018, www.gao.gov/assets/700/690910.pdf.

Tene, Omer, and Jules Polonetsky. "Big Data for All: Privacy and User Control in the Age of Analytics." *Northwest Journal of Technology and Intellectual Property*, vol. 239, 20 Sept. 2012, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2149364.

The Dawn of a New Era: Challenges and Opportunites in Artificial Intelligence. A Alfred Taubman Forum on Public Policy, 12 June 2018, Washington, D.C., Brookings Institution, www.brookings.edu/events/the-dawn-of-a-new-era-opportunities-and-challenges-of-artificial-intelligence/.

The Future Has Begun: Using Artificial Intelligence to Transform Government. Partnership for Public Service and IBM Center for the Business of Government, 2018.

TNW. "A glossary of basic artificial intelligence terms and concepts." Thenextweb.com. https://thenextweb.com/artificial-intelligence/2017/09/10/glossary-basic-artificial-intelligence-terms-concepts/ (accessed 3/22/19).

Tucker, Catherine. "Privacy, Algorithms, and Artificial Intelligence." *National Bureau for Economic Research*, 16 Oct. 2017, www.nber.org/chapters/c14011.pdf.

United States, Congress, *FUTURE of Artificial Intelligence Act of 2017*.

Vardi, Moshe. "Human, Machines, and Work: The Future Is Now." CS 4732. Ethical and Social Issues in AI, 27 Mar. 2017, Ithaca, NY, Cornell University, www.cs.cornell.edu/courses/cs4732/2017sp/.

Wallach, Wendell. *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press, 3 June 2010.

Walton, Robert. "With Artificial Intelligence, It's a Brave New World for Utilities." *Utility Dive*, 24 Nov. 2017, www.utilitydive.com/news/with-artificial-intelligence-its-a-brave-new-world-for-utilities/511008/.

Weinberger, Kilian. "Interpretable Machine Learning: What Are the Limits and Is It Necessary?" CS 4732. Ethical and Social Issues in AI, 13 Mar. 2017, Ithaca, NY, Cornell University, www.cs.cornell.edu/courses/cs4732/2017sp/.

Weld, Dan. "Computational Ethics for AI." CS 4732. Ethical and Social Issues in AI, 20 Mar. 2017, Ithaca, NY, Cornell University, www.cs.cornell.edu/courses/cs4732/2017sp/.

West, Darrell M. "*The Future of Work: Robots, AI, and Automation*" Brookings Institution Press, 2018.

West, Darrell M., and John R. Allen. "How Artificial Intelligence Is Transforming the World." Brookings Institution, *Brookings*, 9 May 2018, www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/.

Wikipedia contributors, "Artificial intelligence," *Wikipedia, The Free Encyclopedia,* https://en.wikipedia.org/w/index.php?title=Artificial_intelligence&oldid=888948971 (accessed March 22, 2019).

Wiseman, Jane, and Stephen Goldsmith. "Ten Great Ways Data Can Make Government Better." Data-Smart City Solutions, Harvard Kennedy School, 11 May 2017, datasmart.ash.harvard.edu/news/article/ten-great-ways-data-can-make-government-better-1041.

Zima, Elizabeth. *Could New York City's AI Transparency Bill Be a Model for the Country?* Government Technology, 4 Jan. 2018, www.govtech.com/policy/Could-New-York-Citys-AI-Transparency-Bill-Be-a-Model-for-the-Country.html.

NATIONAL ACADEMY OF
PUBLIC ADMINISTRATION®

1600 K Street, N.W. Suite 400
Washington, D.C. 20006
Phone: (202) 347 3190
Website: www.napawash.org

ISBN 978-1-7338871-0-6
90000
9 781733 887106