

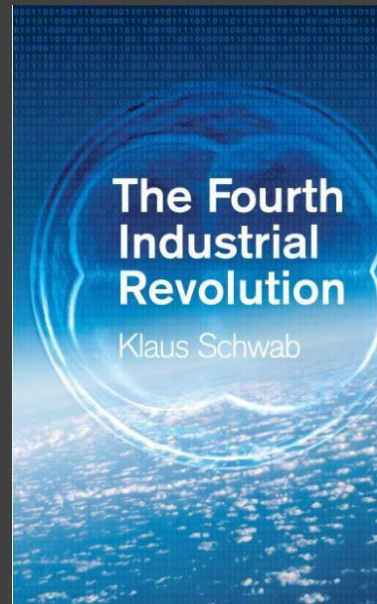
Auditing in the age of artificial intelligence

H. TINA KIM

Today's digital environment is changing the world as we know it. In an era of big data and the digitalization of government records, there is a growing mountain of data available to inform policies and programs. Advances in computing storage and computation skills created a more robust foundation for more sound decisions. However, with this great opportunity comes significant risk.

Klaus Schwab, [The Fourth Industrial Revolution](#)

“The changes are so profound that, from the perspective of human history, there has never been a time of greater promise or potential peril.”



Professor Klaus Schwab, Founder and Executive Chairman of the World Economic Forum, has been at the centre of global affairs for over four decades. He is convinced that we are at the beginning of a revolution that is fundamentally changing the way we live, work and relate to one another, which he explores in his new book, *The Fourth Industrial Revolution*. The world is changing and the pace of change is incredible. Driven by the rise of big data, the commoditization of computing storage, and advances in computing power and tools,

Fourth is evolving at an exponential rather than a linear pace. Moreover, it is disrupting almost every industry in every country. And the breadth and depth of these changes herald the transformation of entire systems of production, management, and governance. The possibilities of billions of people connected by mobile devices, with unprecedented processing power, storage capacity, and access to knowledge, are unlimited. And these possibilities will be multiplied by emerging technology breakthroughs in fields such as artificial intelligence, robotics, the Internet of Things, autonomous vehicles, 3-D printing, nanotechnology, biotechnology, materials science, energy storage, and quantum computing. Already, artificial intelligence is all around us, from self-driving cars and drones to virtual assistants and software that translate or invest. Impressive progress has been made in AI in recent years, driven by exponential increases in computing power and by the availability of vast amounts of data



Amazon scraps secret AI recruiting tool that showed bias against women - Reuters

OCTOBER 9, 2018 / 11:12 PM / A YEAR AGO

Amazon scraps secret AI recruiting tool that showed bias against women

Jeffrey Dastin

SAN FRANCISCO (Reuters) - Amazon.com Inc's (AMZN.O) machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.

As you know AI is fraught with new risks that need to be addressed not the least of which is the potential for amplification of biases that may reinforce existing patterns of discrimination and privacy issues. Just one example in 2018 Reuters reported that Amazon had to scrap its new machine learning recruiting system after finding flaw in its system they had built was showing gender bias. Swimming the recruiting engine was trained to scan through and incoming resumes, observe candidates of past job postings over the previous 10 years since most of the applicants were men still very prominent in Balance across the tech industry, as you know. The recruiting tool developed a bias to prefer male candidates and would penalize resumes that included gender keywords such as women or women's and degraded resumes of those who attended all to all-women's colleges. For everyone and I'm sure that includes everyone gathered here this morning who values and is working to build a diverse and dynamic work workforce. It's critical that we identify and address damaging AI biases like these as well as other risks that you'll be talking about standing squarely on the frontlines and confronting this challenge our accountability professionals who play a key role in helping to ensure the technological advances are implemented responsibly.



Think of candidate selection as a funnel. We have many candidates applying for a job. Out of those, we filter some based on their resumes. Then we continue winnowing the pool at each interview round. Whoever makes it past the last round receives an offer, some of which join the company.

The company's experimental hiring tool used [artificial intelligence](#) to give job candidates scores ranging from one to five stars — much like shoppers rate products on Amazon, some of the people they literally wanted it to be an engine where I'm going to give you 100 resumes, it will spit out the top five, and we'll hire those."

"So it set up a team in Amazon's Edinburgh engineering hub that grew to around a dozen people. Their goal was to develop AI that could rapidly crawl the web and spot candidates worth recruiting, the people familiar with the matter said."

The team had been building computer programs since 2014 to review job applicants' resumes with the aim of mechanizing the search for top talent, five people familiar with the effort told Reuters. [...] In effect, Amazon's system taught itself that male candidates were preferable. It penalized resumes that included the word "women's," as in "women's chess club captain." And it downgraded graduates of two all-women's colleges, according to people familiar with the matter. They did not specify the names of the schools.



The machine does what the human tells it to do. The human instructs, the human chooses the data and the human decides how to act on those outcomes. So there is human interaction at every stage and those are the things that we need to be very intentional about so when you think about that, it's not really artificial.

That is because Amazon's computer models were trained to vet applicants by observing patterns in resumes submitted to the company for over ten years. Most came from men, a reflection of male dominance across the tech industry.

The group created 500 computer models focused on specific job functions and locations. They taught each to recognize some 50,000 terms that showed up on past candidates' resumes.

Secondly, it seems the Amazon team used keywords — 50 000 of them — to score the resumes. That approach is typically known as “bag of words.” All the words found in a resume are thrown together as if the document had no structure.



What is Hireview

based on their facial movements, their tone of voice, their mannerisms

Complaint

Racial Influence on Automated Perceptions of Emotions

Lauren Rhue, University of Maryland - Robert H. Smith School of Business

This study provides evidence that facial recognition software interprets emotions differently based on the person's race.... Both services interpret black players as having more negative emotions than white players; however, there are two different mechanisms. Face consistently interprets black players as angrier than white players, even controlling for their degree of smiling. Microsoft registers contempt instead of anger, and it interprets black players as more contemptuous when their facial expressions are ambiguous. As the players' smile widens, the disparity disappears.

This finding has implications for individuals, organizations, and society, and it contributes to the growing literature of bias and/or disparate impact in AI.

All of this is a remarkably clear-cut illustration of why many tech experts are worried that, rather than remove human biases from important decisions, artificial intelligence will simply automate them. [...] Amazon deserves some credit for realizing its tool had a problem, trying to fix it, and eventually moving on (assuming it didn't have a serious impact on the company's recruiting over the last few years). But, at a time when lots of companies are embracing artificial intelligence for things like hiring, what happened at Amazon really highlights that using such technology without unintended consequences is hard. And if a company like Amazon can't pull it off without problems, it's difficult to imagine that less sophisticated companies can.

Use of AI in Government Hiring

The NYSDOL Skills Matching and Referral Technology (SMART) technology uses artificial intelligence software to analyze résumés for skills and work experience, and identify matches with available employment opportunities. The software compares résumés with openings in New York’s Job Bank and from Job Scout, a tool that synthesizes job postings in other Internet-based job posting and recommends job leads. Utilizing SMART and NYESS, which are already fully integrated, mechanisms can be developed to match individual’s résumés and skills to available job opportunities in state and local government, including available 55-b and -c openings.

Page 13, New York State Employment First Commission

AI is a priority in the federal budget. There are multiple executive orders. first and foremost emphasizes the need for the US to improve its research and development in AI. It states that the US must “drive technological breakthroughs in AI” to “promote scientific discovery, economic competitiveness, and national security.”

Trump’s AI plan asks federal agencies to prioritize research in the field by reallocating resources within their existing budgets, calling on the heads of federal agencies to “consider appropriate administrative actions” to do so.

Both our defense side. And in Congress, we have elevated the conversation around AI and I want to point to something on the bottom right there that is of personal excitement to me the federal data strategy the first for the first time we have a federal data strategy. And for those of you who are watching virtually strategy.data.gov has that information and that's our commitment to improving the quality of data the availability of data, but ensuring that we're doing it in a way that protects privacy and security. It preserves the intent of use of data, so we collected it and that we inform citizens why we collected and what we intend to do with it and the protocols around ethics. You heard that Chief data officer have been aligned across the entire federal government and that's part of the discipline that we're trying to drive because use of the powerful tool the foundation is data and that's where we're spending a significant amount of time now inside federal agencies. We're making real progress in R D and for the full spectrum. These are a couple of examples right that's those are the framework for what we are intending to do but here's some examples when we think inside the federal government what those numbers that Tina shared what are they really doing?

Use of AI in Government Hiring

The NYSDOL Skills Matching and Referral Technology (SMART) technology uses artificial intelligence software to analyze résumés for skills and work experience, and identify matches with available employment opportunities. The software compares résumés with openings in New York's Job Bank and from Job Scout, a tool that synthesizes job postings in other Internet-based job posting and recommends job leads. Utilizing SMART and NYESS, which are already fully integrated, mechanisms can be developed to match individual's résumés and skills to available job opportunities in state and local government, including available 55-b and -c openings.

Page 13, New York State Employment First Commission

They kind of go into three buckets. They're using AI and data to serve agency mission, to accomplish what the agency is intended to do, to save citizens, food water, first responders, our military, financial fidelity, many of the things in that space what you see though there is the areas that we're working on Workforce Technology hardware and infrastructure and operational integration.

There are a lot of Pilots going on, but now those great results have to go somewhere and we have to make those part of the way we deliver business federal agencies are also responsible for stimulating the research community that is a whole separate, you know piece that I showed you on the first page, so that means grants but it also means making Data available externally It is part of the responsibility of the federal government to make data available to stimulate the research and development that is going on in our country. And in fact in the federal data strategy, there were three priority data sets. And when I say that that means where we have asked agencies to apply intensity and some common approaches to make those more available. Financial management data geospatial data and data that has been requested by the industry to support AI research and development that's an important area. And then the last piece is American industry. So there are when you talk about autonomous vehicles when you talk about new products and services for Citizens, there are many things that to stimulate our economy and grow our businesses in the United States. They need data to do that and some of the very important data is inside the federal government, but we have to do that in a way you see those pieces at the bottom. And that's a very important consideration for this community. We have to do that in a way that aligns with regulations standards guidance and the intent for which we have collected that data from citizens. So we know that as we think about. These capabilities inside the federal government the nature of the services that we can deliver to Citizens change every one of you think about ask you to think about your personal experiences, right you probably if you got an Uber to get here, you're probably using you know, you have that you were experiencing AI if you play video games with your kids or your grandkids and they probably beat you have seen those capabilities applied. You experience these things in your daily life.

Illinois [HB 2557](#), The Artificial Intelligence Video Interview Act

Provide notice: Employers must inform applicants that AI will be used to analyze their interview videos. (A provision requiring written notice was removed from the bill.)

Provide an explanation: Employers must explain to the applicant how their artificial intelligence program works and what characteristics the AI uses to evaluate an applicant's fitness for the position.

Obtain consent: Employers must obtain the applicant's consent to be evaluated by AI before the video interview and may not use AI to evaluate a video interview without consent.

Maintain confidentiality: Employers will be permitted to share the videos only with persons whose expertise or technology is needed to evaluate the applicant.

Destroy copies: Employers must destroy both the video and all copies within 30 days after an applicant requests such destruction (and instruct any other persons who have copies of the video to destroy their copies as well).

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica May 23, 2016

...computer program spat out a score predicting the likelihood of each committing a future crime. Borden — who is black — was rated a high risk. Prater — who is white — was rated a low risk.

Two years later, we know the computer algorithm got it exactly backward. Borden has not been charged with any new crimes. Prater is serving an eight-year prison term for subsequently breaking into a warehouse and stealing thousands of dollars' worth of electronics.

Scores like this — known as risk assessments — are increasingly common in courtrooms across the nation. They are used to inform decisions about who can be set free at every stage of the criminal justice system, from assigning bond amounts — as is the case in Fort Lauderdale — to even more fundamental decisions about defendants' freedom. In Arizona, Colorado, Delaware, Kentucky, Louisiana, Oklahoma, Virginia, Washington and Wisconsin, the results of such assessments are given to judges during criminal sentencing.

US STATES USE OF AI: 2019 NASCIO Survey

Currently Using	13%
Piloting	24%
Proofs of concept/demonstrations	31%
Evaluating/gathering requirements (RFI)	19%
No use or planned use	13%

So why is this a concern? This is the results of a recently released survey of 45 state CIOs which shows level of use of AI in State government-while state CIOs are moving cautiously into AI only 13% report no use or planned use. The rest are either using AI, piloting AI or in the process of evaluating or performing proofs of concepts. Moreover, Thirty-two percent of survey respondents said departments or agencies are deploying AI without the CIOs involvement and another 27 percent say they don't know if their agencies were doing so. Additionally, 59 percent of the respondents to this survey did not currently have a AI framework to evaluate risk, predictive policing, pretrial risk assessments, school-assignment matching algorithms, fraud-detection systems, traffic-management systems, screening tools, or face recognition.

New York City: Local Law 49

This law required the creation of a task force that provides recommendations on how information on agency automated decision systems may be shared with the public and how agencies may address instances where people are harmed by agency automated decision systems.

LL49 requires that the Mayor establish and convene a task force to examine how the City uses ADS in decision-making, and provide specific recommendations related to the following: Criteria for identifying which agency automated decision systems should be subject to one or more of the procedures recommended by such task force...; Development and implementation of a procedure through which a person affected by a decision concerning a rule, policy or action implemented by the City, where such decision was made by or with the assistance of an agency automated decision system, may request and receive an explanation of such decision and the basis therefor; Development and implementation of a procedure that may be used by the City to determine whether an agency automated decision system disproportionately impacts persons based upon age, race, creed, color, religion, national origin, gender, disability, marital status, partnership status, caregiver status, sexual orientation, alienage or citizenship status; Development and implementation of a procedure for addressing instances in which a person is harmed by an agency automated decision system if any such system is found to disproportionately impact persons based upon a category described in [the above paragraph]; Development and implementation of a process for making information publicly available that, for each agency automated decision system, will allow the public to meaningfully assess how such system functions and is used by the City, including making technical information about such system publicly available where appropriate; and, The feasibility of the development and implementation of a procedure for archiving agency automated decision systems, data used to determine predictive relationships among data for such systems and input data for such systems, provided that this need not include agency automated decision systems that ceased being used by the City before the effective date of this local law.

NYC AI Task Force

Part of the challenge for task force members was that city agencies, accustomed to keeping details of vendor technologies secret, [were unwilling to provide even a basic list of automated systems already in use.](#)

“There were delays, and obfuscations, and then, by the spring of 2019, outright denials. It wasn’t because the data didn’t exist,” wrote Albert Fox Cahn, executive director of civil rights and privacy group Surveillance Technology Oversight Project, who participated in some task force meetings, in FastCompany. “The City Council passed the task force into law to hold the administration accountable for its secretive use of algorithms, but that was the last thing the administration wanted.”

Survey Logistics

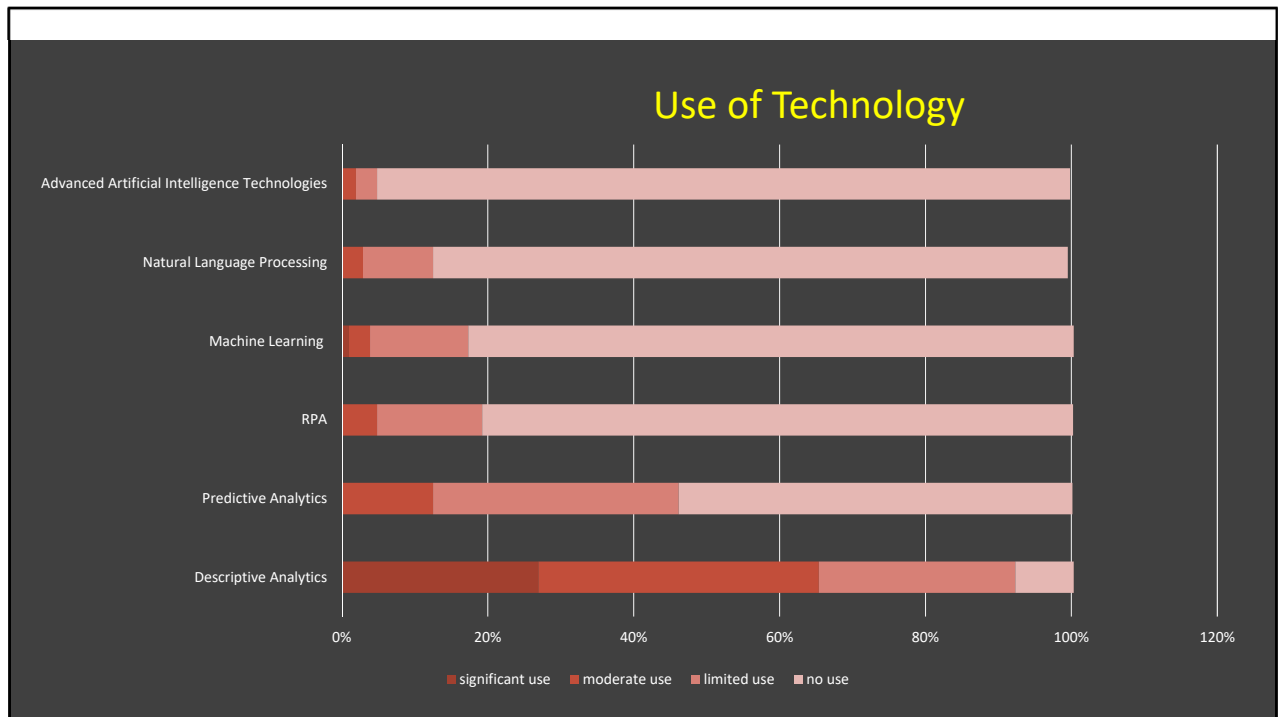
We administered surveys during July 2019 to members of the following organizations using Survey Monkey:

- Council of the Inspectors General on Integrity and Efficiency (CIGIE)
- National State Auditors Association (NSAA)
- Association of Local Government Auditors (ALGA)

The survey was administered using survey monkey to federal Inspector Generals, state auditors, and local government auditors.

There are many different definitions of AI. So in order TO **provide a standard frame of reference** WE used definitions employed by the Institute of Internal Auditors and their AI framework

We received 125 responses out of a populations of over 400 but only 105 included demographic information-so the response rate of usable responses was approximately 25%.



Over 90% of agencies reported some use of analytics but their use of other technologies was significantly lower

46% reported some use of predictive analytics.

19% reported some use of Robotic Process Automation

19% reported some use Machine Learning

13% report some use of Natural Language Processing and

5% report some use of Advanced AI (deep learning and neural networks)

These results are similar to results reported by the IIA in 2019 related to public sector internal auditors-However, other than descriptive analytics the level of use was consistently lower for our survey what the IIA reported.

For instance for Machine Learning, which was the highest use category for AI, only 1 respondent indicated significant use and only 3 indicated moderate -use. All other respondents indicated limited use and this is likely an overestimation. Of the 15 respondents (primarily locals) who indicated limited use there appeared to be confusion regarding what constituted AI work and their detailed answers did not seem consistent with the answer provided (for example, replying that they were using a tool that is not commonly associated with AI work or more commonly reporting that were not using any tools at all.)

The Role of Audit in Artificial Intelligence

From The IIA's AI Framework


For organizations that use AI, Internal Audit should consider AI as part of its risk assessment and consider whether to include AI in its risk-based audit plan.

For organizations exploring AI, internal audit should be actively involved in AI projects from their beginnings, providing advice and insight contributing to successful implementation.

For organizations that have implemented some aspect of AI, internal audit should provide assurance on management of risks related to the reliability of the underlying algorithms and the data on which the algorithms are based.

Internal audit should ensure the moral and ethical issues that may surround the organization's use of AI are being addressed.

Like the use of any other major system, proper governance structures need to be established and internal audit can provide assurance in this space.



The IIA's AI Framework

The Framework is comprised of six components, all set within the context of an organization's AI strategy.

Components:

- AI Governance
- Data Architecture and Infrastructure
- Data Quality
- Measuring Performance of AI
- Human Factor
- Black Box Factor

Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy



Argued that poorly rendered algorithms reinforce discrimination and widen inequality.



Solution: Algorithmic audit.



Human Factor

Relevant Ethics Objectives and Activities or Procedures

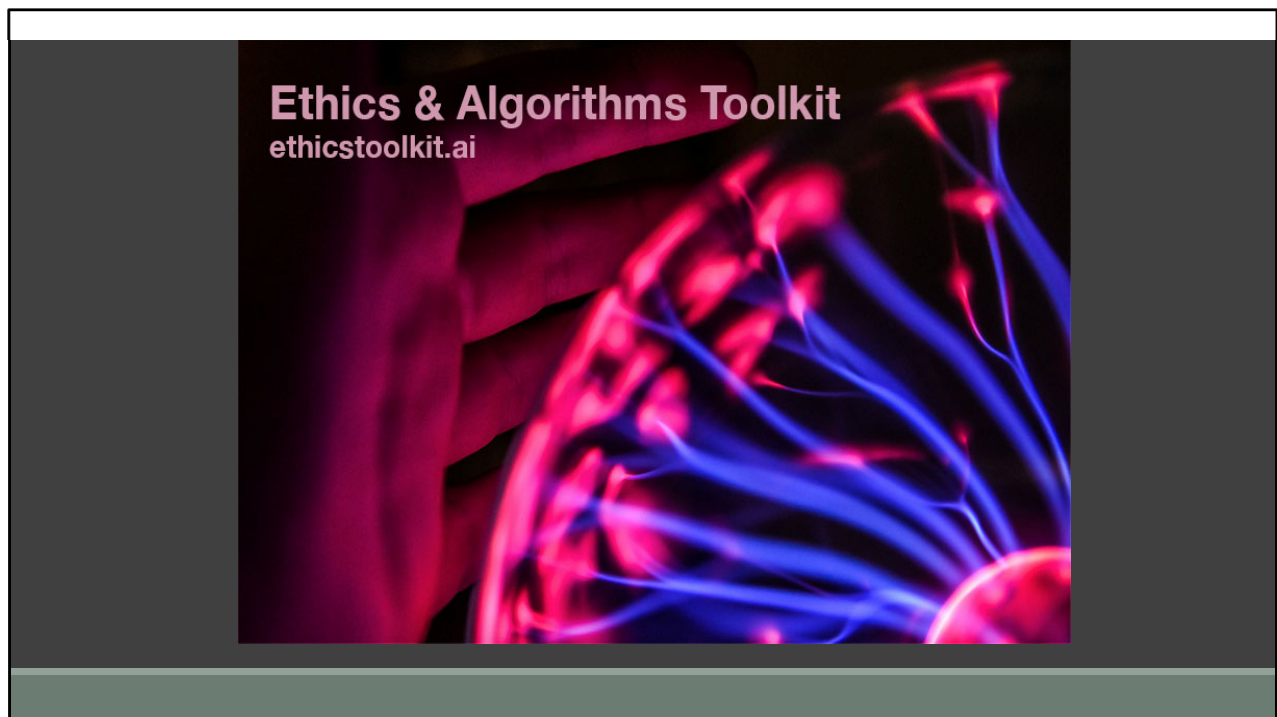
Engagement or Control Objective(s): Provide assurance that outcomes of the organization's AI activities are free from unintended biases.

Activities or Procedures: Review the intended results of the AI activities (strategic objectives) and compare with actual results. If a variance is detected, determine if bias is the cause.

IEEE (Ethically Aligned Design)

1. **Human Rights**—A/IS shall be created and operated to respect, promote, and protect internationally recognized human rights.
2. **Well-being**—A/IS creators shall adopt increased human well-being as a primary success criterion for development.
3. **Data Agency**—A/IS creators shall empower individuals with the ability to access and securely share their data, to maintain people's capacity to have control over their identity.
4. **Effectiveness**—A/IS creators and operators shall provide evidence of the effectiveness and fitness for purpose of A/IS.
5. **Transparency**—The basis of a particular A/IS decision should always be discoverable.
6. **Accountability**—A/IS shall be created and operated to provide an unambiguous rationale for all decisions made.
7. **Awareness of Misuse**—A/IS creators shall guard against all potential misuses and risks of A/IS in operation.
8. **Competence**—A/IS creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation.





GovEx, the City and County of San Francisco, Harvard DataSmart, and Data Community DC have collaborated on a practical toolkit for cities to use to help them understand the implications of using an algorithm, clearly articulate the potential risks, and identify ways to mitigate them.

There are five documents which comprise the toolkit:

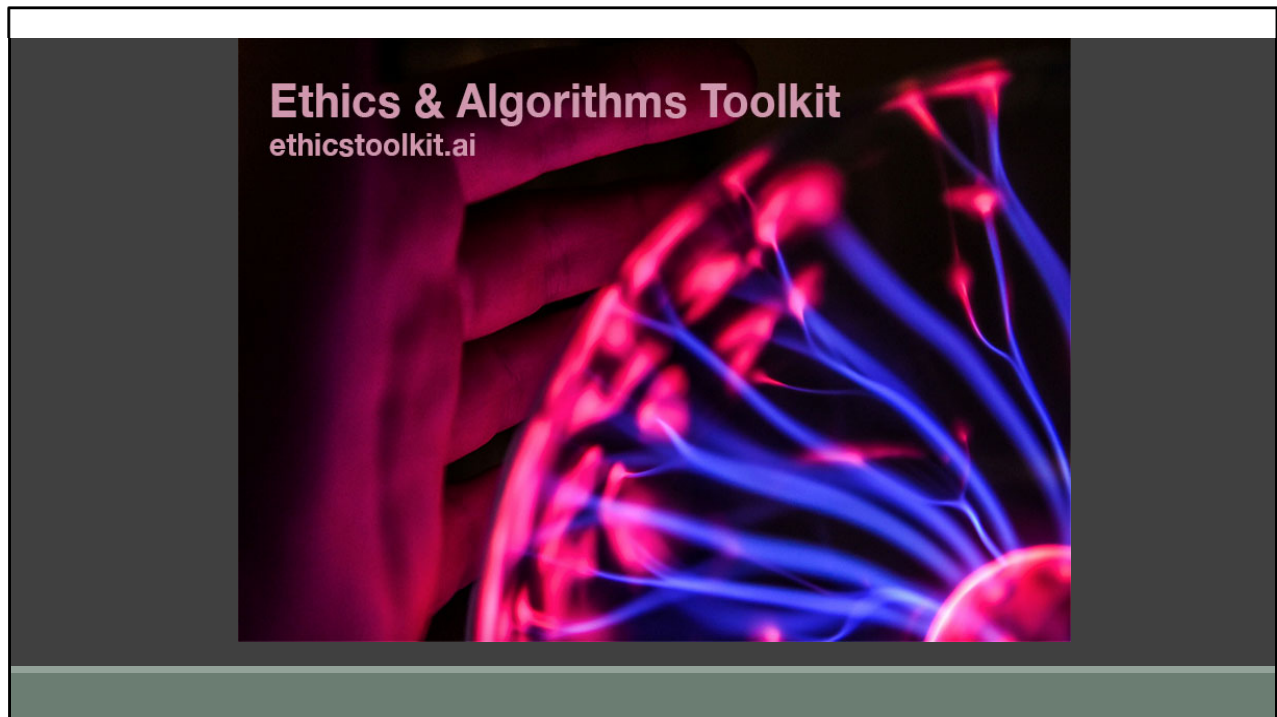
Overview and Introduction

The [overview section](#) of the toolkit is comprised of level-setting background information that will be useful when traversing subsequent sections of the toolkit. We have outlined a few real-life scenarios where the toolkit might be applied, provided definitions, and more. For example, while we briefly touched upon machine learning in the previous module, the toolkit overview helps you understand more about the various types which exist, such as supervised learning, unsupervised learning, and so on.

Part 1: Assess Algorithm Risk

In [Part 1](#) of the toolkit, there are six major steps (or questions) to help you and your stakeholders characterize an algorithm. Many of these steps have multiple components, but also include clear instructions on how to summarize those stages in order to complete the step.

Since this document can be difficult to navigate, we have developed a [worksheet for Part 1](#), designed to help you track your responses to the individual steps and how they are combined into overall risk values. It's worth noting that although answering a series of questions seems simple, you will almost certainly need additional people to help - whether they are stakeholders, data analysts, information technology professionals, or representatives from a vendor that you are working with. Don't expect to complete this part of the toolkit in just a few hours. Some of the steps will evoke considerable discussion.



Part 2: Manage Algorithm Risk

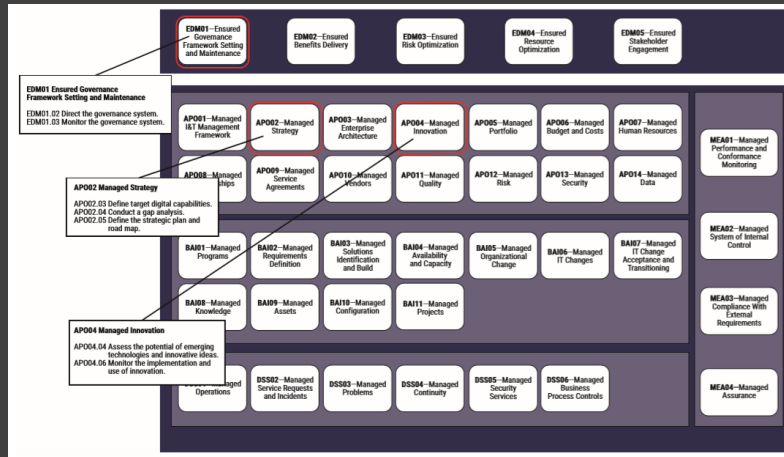
Although it's helpful to know how concerned you should be about various aspects of your algorithm, that's really only half the battle. Although there may be a few cases where the risks are too severe to proceed, there are often ways to mitigate them. Using [Part 2](#) of the toolkit, you identify specific techniques to help address the considerations you identified in Part 1.

The results of Part 2 will be highly customized and specific to the factors you evaluated in part 1. Some of the recommendations can introduce significant burdens that are more appropriately addressed within large-scale programs, such as those that support the social safety net. It is not unusual to need executive and political support to be successful.

Appendices

Although this isn't specifically required reading in order to use the toolkit, the [appendices](#) provide plenty of additional context and depth. The first appendix contains a list of in-depth questions to help you understand your data in more detail. The second provides additional background on bias and how easily it can arise.

Relevant COBIT® 2019 Governance and Management Objectives



24

*available on isaca.org

Figure 1 highlights several examples of processes within COBIT® 2019 that may provide help in compiling a list of risks and controls for the AI initiative within an organization.

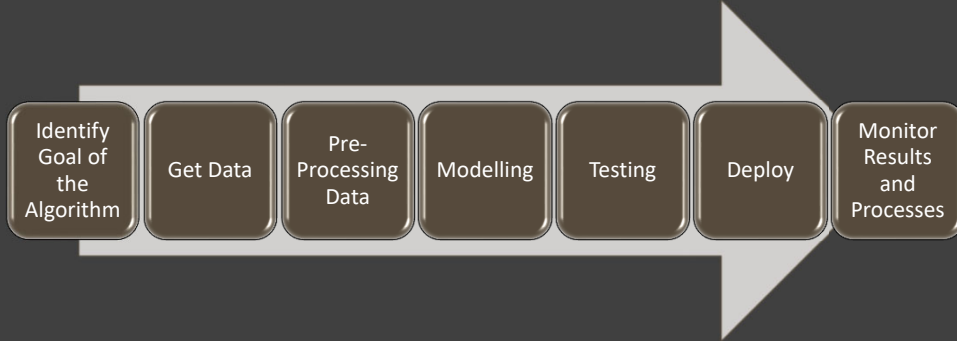
Auditing AI

What would it cover?

What can and cannot be audited?

What Would an AI Audit cover?

Identify audit elements through the general AI processing pipeline:



Questions

H. Tina Kim

Deputy Comptroller for State Government Accountability

Office of the New York State Comptroller

tkim@osc.ny.gov