NATIONAL ARCHIVES

OFFICE *of the* CHIEF RECORDS OFFICER

Cognitive Technologies White Paper

Records Management Implications

for

Internet of Things, Robotic Process Automation, Machine Learning, and Artificial Intelligence

National Archives and Records Administration

October 2020

Abstract

The National Archives and Records Administration (NARA) performed a review and analysis of Internet of Things (IoT), Robotic Process Automation (RPA), Machine Learning (ML), and Artificial Intelligence (AI), collectively referred to in this paper as cognitive technologies. The goal was to provide forward-thinking perspectives on these technologies and their potential impact on Federal records management.

This white paper provides a basic description for each of the cognitive technologies, examples of their various applications, and several enabling factors affecting the technologies. It focuses on common algorithm training methods and how the programming advances from providing basic outputs and functions into more intuitive or AI enabled systems. AI may provide nuances in human-like decision-making and this white paper explains biases and ethical concerns related to the development and training of algorithms.

The paper also provides an analysis of records management implications of cognitive technologies. The analysis starts with the correlation between records and data, specifically highlighting both are subject to similar governance structures and NARA requirements. It focuses on policy and standards along with risks associated with the authenticity and integrity of records. Additionally, the paper addresses the implications surrounding records appraisal, scheduling, and transfer.

This white paper also describes specific challenges around data because of its exponential growth rate and use. As new ways of creating and capturing data are created, the management of data often lags behind. While the volume can be staggering, the data has to be managed by agencies within the appropriate records management or data framework.

Introduction

Human decision-making is being supported or replaced by a number of emerging technologies. The National Archives and Records Administration (NARA) performed a review of four specific emerging technologies: Internet of Things (IoT), Robotic Process Automation (RPA), Machine Learning (ML), and Artificial Intelligence (AI), which are referred to as cognitive technologies.

This white paper describes these four cognitive technologies and provides various use cases and examples of how records and data are created, including government records and data. The paper also describes several enabling factors impacting the application of the cognitive technologies. The paper examines how their capabilities can be amplified beyond their individual purposes when coupled with one another. In addition, the paper also discusses AI decision-making as well as potential biases and ethical concerns.

As records and archives professionals, there is ongoing interest to learn if these technologies have advanced to the point where they can seamlessly automate into the management of Federal records. The paper analyzes and provides forward-thinking perspectives on how cognitive technologies could automate Federal records management functions.

The paper also describes existing standards and policies that impact the creation and management of Federal records and data. There is a strong correlation between records and data: both are subject to similar governance structures and NARA requirements. The exponential growth of data creation and data collection resulting from cognitive technologies requires forethought on how to adapt and develop standards and policies to ensure that agencies and NARA can meet the records and information management challenges and take advantage of the opportunities presented by cognitive technologies.

Internet of Things

The term Internet of Things (IoT) was first used by British technologist Kevin Ashton in 1999 to argue for connecting Radio Frequency Identification to the internet. Conceptually, IoT has grown to encompass almost any device that has a microprocessor and can communicate wirelessly.

> *The Internet of Things, or IoT, refers to the billions of physical devices around the world that are now connected to the internet, collecting and sharing data. Thanks to cheap processors and wireless networks, it's possible to turn anything, from a pill to an aeroplane to a self-driving car into part of the IoT. This adds a level of*

digital intelligence to devices that would be otherwise dumb, enabling them to communicate real-time data without a human being involved, effectively merging the digital and physical worlds (Ranger, 2018).

Applications

IoT devices use sensors to collect data and they are being deployed in almost every facet of human activity. Applications of IoT devices can be found in our personal lives, as well as in industrial and governmental sectors. Examples include:

Personal

Smart devices, such as virtual assistants, have become increasingly common in people's lives. Smart assistants are programmed to recognize voice commands and are able to provide information and perform simple tasks. Smart assistants can be used in conjunction with home security systems to operate smart home devices, such as lights, door locks, and window blinds or curtains. They can perform administrative functions, such as add calendar items, create lists and reminders, answer questions, and dial phone contacts. They can be used recreationally to turn on a fireplace, play select genres of music, and tell jokes. Many smart devices can be paired, interconnected, and remotely accessed and controlled to automate homes, also commonly known as smart homes.

Industrial

In addition to the inclusion of IoT devices in our personal lives, IoT devices have industrial and utilitarian uses in many areas, including automotive and aviation services. For example, in 2018, Rolls-Royce rolled out its IntelligentEngine, which uses IoT sensors to capture 70 trillion data points for tracking aircraft engine maintenance and repairs (Rolls-Royce, 2018). The system is network-connected and uses data points to automatically suggest maintenance.

Government

A city that uses IoT technology to collect data to manage resources and services is known as a smart city. In 2016, the City of San Diego approved a plan to retrofit streetlights with processors and data storage; Bluetooth and Wi-Fi radios; 1080p video cameras; acoustical, temperature, pressure, humidity, vibration, and magnetic fields sensors. The city deployed the IoT devices to gather data and analytics to improve parking, traffic, and safety (Potter, 2019). Despite the potential benefits, San Diego's use of IoT devices in public areas has highlighted awareness about privacy, retention, and access to the data.

In 2017, the United States Geological Survey (USGS) rolled out an earthquake early warning system, ShakeAlert. The system is being co-developed by universities, public and private partners, and its technology is based on a network of sensors designed to detect and communicate real-time seismic information to a centralized control center. ShakeAlert will provide seconds, minutes, and sometimes hours to allow people life-saving moments to take safety or protective measures (U.S. Geological Survey, 2017).

Enabling Factors

Some experts estimate that by 2025 there will be 75.44 billion devices connected to the internet world-wide (Statista, 2016). Another study estimates that by 2025 nearly 30% of all data created, captured or replicated will be created in real-time (Reinsel, Gantz, Rydning, 2018). The following five factors are contributing to the growth of IoT.

Internet Version Protocol 6

An Internet Protocol (IP) is the network address for a device connected to the internet that allows data to be sent to and from a device. Until recently, an IP address was based on a 32-bit protocol that allowed for approximately 4.3 billion unique IP addresses: four groupings of numbers ranging from 0-255 (e.g. 128.105.39.11). The rapid growth of the internet made it clear that the total number of devices would eventually outstrip the protocol's capacity. A new IP version, IPv6, was ratified in 2017 and specifies eight groupings of 4 hexadecimal digits (e.g. 2a03:2880:2110:df07:face:b00c::1). This allows for 3.4×10^{38} number of IP addresses or enough to allow unlimited growth of IoT devices.

Drop in Sensor Price

Data from Goldman Sachs indicates the average cost of a sensor in 2018 was $.44 per sensor, which was nearly 200% less than the average cost in 2004 (Microsoft, 2018). Decreasing cost combined with ease of deployment, and Wi-Fi connectivity dramatically lowers the barriers organizations face when deploying sensors, or IoT devices, across their business processes.

Edge Computing

Sometimes when users request data from the cloud, there is a delay before the data is returned. This is known as latency. If a user is searching online for a restaurant, latency can be annoying, but not important. Latency in multi-player online games could make them unplayable. Latency in an autonomous driving truck could be dangerous, even fatal. Data processing that can happen locally will likely be faster than sending and receiving data from the cloud. Edge computing is the effort to push data processing away from the centralized cloud to the physical location or edge where the data is needed to speed response (Miller, 2018). When data processing happens locally, it is faster than sending and receiving data from the cloud. By embedding the AI sensor in IoT devices, as opposed to some cloud platform hundreds or thousands of miles away, edge computing will reduce latency for IoT technologies.

Network Improvements

Receiving and transmitting data is critical for IoT devices. Improvements in networks such as 5G and gigabit fiber ensure high upload and download speeds are available with very low latency. For example, 4G data has to be transmitted between 100Mbs and 1Gbs. Theoretically, 5G will achieve 1-10Gbs. Aside from data transmission speed, the biggest impact will be the drop in latency: 4G has a latency as low as 50 milliseconds, but 5G is intended to have latency less than 1 millisecond.

Geographic Information Systems

Geographic information systems is an umbrella term for technologies that provide location services. These services are based on determining the spatial location of a sensor on the earth often captured as latitude and longitude and perhaps altitude. Commonly, geographic information system data is coupled with other metadata, such as temperature or speed, to help visualize or provide context.

Collectively, these five enabling factors assist the availability, increased capacity, and data transmission speed of IoT.

Robotic Process Automation

Robotic Process Automation (RPA) is a technology platform that enables a software robot to interact with applications (The American Council for Technology-Industry Advisory Council, 2019). More clearly, this means RPA is software code that instructs a device or electronic system to perform a function. The device or system is referred to as a bot.

In order for an RPA to execute the instructions in the coding, or learn a process, a program engineer must create a detailed task list and “teach” these steps to the bot. This means the engineer must write a program script for the bot to follow and execute its tasks repeatedly and quickly. Structured input, rule-based processes, and structured output is required for an RPA to optimally perform. The dependence on a detailed task list or a script means that an RPA cannot easily process deviations, such as errors or wrong information.

Applications

Email and Website Interface

RPA software is designed to reduce the burden of repetitive and simple tasks on employees. For example, a member of the public could send an email to NARA's reference desk asking for information about a topic. An RPA system could open the email, scan the content, open the reference request system, populate the requester's contact information, and make a best guess at the purpose of the request. With this software being virtual, an organization can create as many instances of an RPA as required to meet demand. It has become common on websites to offer users the option of interacting with a chat feature to find information faster. In 2018, NARA launched its own chatbot to help answer questions from the public and reach the agency's strategic goal "Connect with Customers" (Wright, 2018).

Digitization, Data Entry, and Business Process Automation

Some of the early proponents of RPA were companies offering digitization software. Digitization software has focused heavily on reading images through optical character recognition, extracting data from forms, and delivering structured data into databases. Companies began expanding their digitization business portfolios by using RPAs to perform data entry, which in turn enabled them to market their services to include business process automation. RPAs offer the opportunity to automate multiple steps in a business function.

RPA is beginning to impact business processes across multiple sectors. Common themes in successful implementations focus on increased data entry accuracy and reduced staffing costs. For example, a bank deployed 85 bots to run 13 processes, handling 1.5 million requests per year. RPA allowed the bank to add the capacity equivalent to more than 200 full-time employees at approximately 30 percent of the cost of recruiting more staff (Schatsky *et al*, 2016).

Enabling Factors

RPA is a form of business process automation that uses the Graphical User Interface (GUI) of a system to automate actions. RPAs interface with the front end, or GUI, of an application as a human would and can therefore imitate human data entry. Generally, an RPA is focused on automating straightforward tasks like answering simple queries, opening applications, and keying rather than interpreting data. Multiple RPAs can be combined and this process of overlaying allows the system to make decisions regarding information it processes.

Virtual Interface

The GUI was initially designed in response to the challenges of using command line interfaces required of early computer programs. General users preferred icons rather than typing commands. RPA takes this concept one step further by allowing a computer to use a virtual keyboard and mouse to interface with a GUI. RPA uses the same type of signals as a mouse or keyboard to connect with software, but does so virtually. Through virtualization, multiple bots can be used to key data into a software GUI simultaneously.

Cost, Agility, and Benefits

Sometimes proprietary programming knowledge creates a barrier to broader automation. RPA is quicker and easier to program because it uses the GUI rather than writing code. This means RPAs can be created more cheaply and deployed faster than writing code. Once an RPA is implemented, updates and maintenance has the potential to impact the bot or render it useless if there is any change to the GUI or business process.

The use of RPAs is pertinent to agencies interested in digitizing their records in preparation for meeting requirements in the joint Office of Management and Budget (OMB) and NARA Memorandum, M-19-21, Transition to Electronic Records. The memorandum requires agencies to manage temporary and permanent records electronically and with the appropriate metadata. It also encourages digitization because it requires agencies to close agency-operated records storage facilities. It is now common for service providers to leverage RPAs for their digitization services, which means agencies should be carefully assessing market research and fair market value to ensure judicious government spending.

One argument for RPA is that it frees up human workers from mundane tasks and allows them to focus on more creative or challenging tasks. Other benefits RPAs provide is the ability to evaluate outputs from tasks or processes completed by humans to ensure consistency and accuracy. According to Leslie Willcock, Professor of Technology Work and Globalisation at the London School of Economics and Political Science, we have been trying to turn humans into robots, but now with RPA we can “take the robot out of the human” (Lhuer, 2016). While true, RPA is also reducing the need for organizations to hire staff to perform repetitive skill tasks. In anticipation of the Federal Government implementing more RPAs into their work, the Office of Personnel Management issued the Reskilling Toolkit in 2019 to help agencies prepare for this shift in resources (Office of Personnel Management, 2019).

Machine Learning and Artificial Intelligence

Algorithms are fundamental to both machine learning (ML) and artificial intelligence (AI). Dictionaries define an algorithm as a step-by-step procedure that frequently involves repetition, for solving a problem or accomplishing some end. An example of an algorithm could be a recipe for ice cream: four ingredients mixed and frozen properly will create a refreshing dessert.

The term machine learning (ML) refers to a software programming technique that uses algorithms to autonomously improve decisions through analysis. The algorithms use statistical methods that enable machines to improve correlations as more data is used. This facilitates the machine's ability to automatically discover patterns in data which can be used to make predictions. The algorithms generally perform better as the volume of data available to analyze increases (Thompson).

Artificial Intelligence (AI) can be described as teaching machines to learn and solve problems so they can make yes or no decisions (Sibley, 2018). Algorithms, methods, or technologies are employed to make a system behave like a human. AI algorithms are already used in our daily lives, from Netflix predicting what we want to watch next to a job application system that determines if we are likely candidates for a position. AI can be classified as a lower or higher stake system depending on the impact. Lower stake systems, such as Netflix, predict what we would want to watch next. AI shapes what we want to watch based on people's searches. Higher stake systems use data to make decisions and predictions about people. These high stake AI systems have real world implications, meaning they influence life more than a bad movie selection on Netflix. An example of a higher stake system is the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) system used to predict whether an offender is likely to recommit a crime (Broad, Recordkeeping Roundcasts, 2018).

AI makes use of multiple learning techniques, including ML, natural language processing, sensors, and human-computer interactions. ML fits within the broader concept of AI.

Applications

Email Management

An example of machine learning for email management, research, and e-discovery is Maura Grossman and Gordon Cormack partnership with the Library of Virginia to use ML tools to classify former Virginia Governor Timothy Kaine's emails. The system used a protocol, Continuous Active Learning (CAL), with a technology-assisted review (TAR) tool to review and classify the governor's email records for public release. CAL categorizes emails by relevant subject within large data sets and allows the ability to release open information, and separate and withhold disclosure of restricted materials. Grossman and Cormack conducted an experiment with the Kaine emails and worked with the Library to launch a website that enables the public to search and more easily access the governor's emails. The Library is the first state archives to achieve this accomplishment in promoting transparency (COSA Study 5, 2017).

In Australia, there was a similar project to use ML for email appraisal at the Public Record Office Victoria (PROV). The PROV had 67,000 tapes and 28 petabytes of content and could no longer efficiently provide access to the records. The project explored using a TAR tool to conduct technical appraisal and de-duplicate a sample of the emails. While the tool was effective, it was not completely automated. In addition to the tool, it was necessary to have traditional appraisal knowledge to construct the searches (Rolan, 2019).

Digitization and Metadata

A recent proof-of-concept study at the National Herbarium of the National Museum of Natural History, Smithsonian Institution at Washington, D.C. explored using deep learning to detect mercury-stained botanical specimens. The botanists manually created two datasets of mass digitized images, a set of 7,777 unstained specimens and 7,777 stained images which were used for training and testing. The researchers used the algorithm to identify two types of plants, 9,276 clubmoss and 9,113 spikemoss. They layered the data to identify species and differentiate between unstained and stained. The success rate was over 90%. The researchers emphasized that this work requires proper metadata and many hours visually identifying the specimens for the original training set (Schuettpelz, 2017). Programs such as this may handle preliminary specimen categorization and could facilitate conducting comparisons of digitized images. Conducting this project required interdisciplinary skills and knowledge of data scientists, digitization experts, and botanists (Smith, 2017).

Recordkeeping

James Lappin, in his Thinking Records blog, envisions AI's impact on recordkeeping with algorithms that will assign access and retention rules. Instead of relying strictly on metadata, the algorithm will be able to identify patterns within any type of data. While these algorithms will generate data, they will also be used to manage it. The challenge, according to Lappin, will be for the profession to use AI functionality. He raises the possibility of AI restructuring an entire records system to enable the application of access and retention rules to a different set of aggregations from when the records were created. (Lappin, 2020)

Enabling Factors

Algorithms are one of the primary components that comprises ML and AI. This section explores the different methods that can be used to refine algorithms and improve outputs.

Algorithm Training Methods

The data or outputs of ML and AI systems are only as good as its programming or training methods. Numerous learning models are used to validate the predictive information or decision-making machine learning provides. This paper reviews three training methods: Supervised, Unsupervised, and Reinforcement (van Essen, 2019). Figure 1, Learning Methods for Machine Learning and Automatic Classification, depicts characteristics of each learning method: the task driven supervised method, the data driven unsupervised method, and the learning from mistakes reinforcement method.

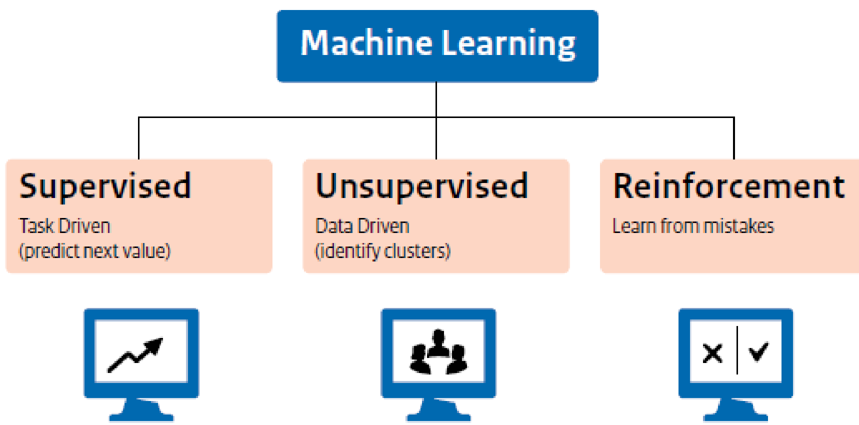

Figure 1. Learning Methods for Machine Learning and Automatic Classification, The Nationaal Archief, 2018

Supervised Learning

The supervised training method works by using labeled elements or image processing to classify data, such as cars, people, or birds. This method requires a human to manually label input information or data. The algorithm is then trained to make accurate decisions to predict the labeled data, such as determining if a digital object represents a car, a person, or a bird. This training method is manual and time consuming, which makes it very expensive. The initial introduction of clean data to the algorithm yields more accurate results or predicted outputs.

There are commercial services, such as [Amazon Mechanical Turk] (MTurk), that leverage crowdsourcing as way organizations can purchase services to provide data entry or labeling. This service in some cases may off-set the expense involved with the supervised training method.

Unsupervised Learning

Unlike supervised learning, the unsupervised training method does not require labeled data. Instead, the data is formatted to pre-set specifications and the algorithm is trained with the data. The algorithm identifies similarities in data and then groups it together in clusters for further analysis as shown in Figures 2 and 3. While this model provides more accurate outputs there is no transparency with how the algorithm came up with similarities in the data clusters.

The process of clustering similar groups of data to provide accurate system outputs leads into predictive analysis. The outputs are considered predictive because they mimic a human's ability to learn and anticipate.

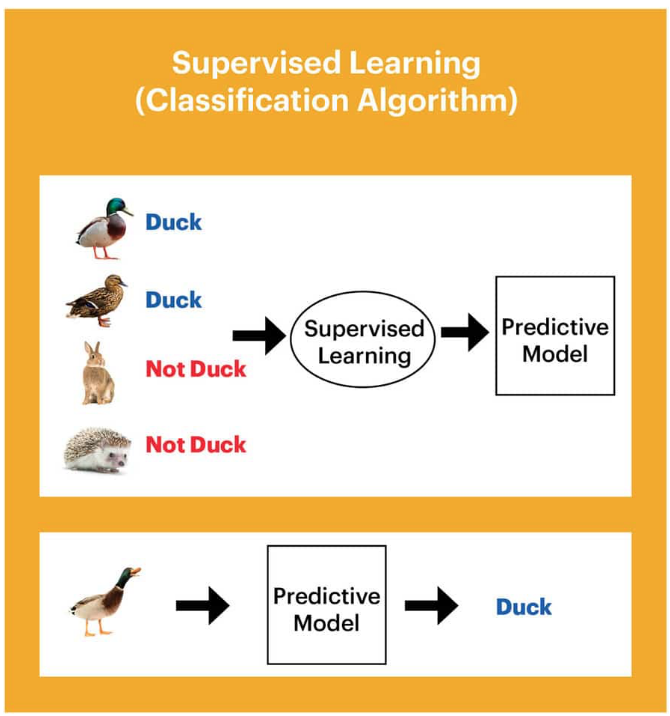

Figure 2, Supervised learning using subject matter experts to "teach" the correct choices (Zhou, 2018)

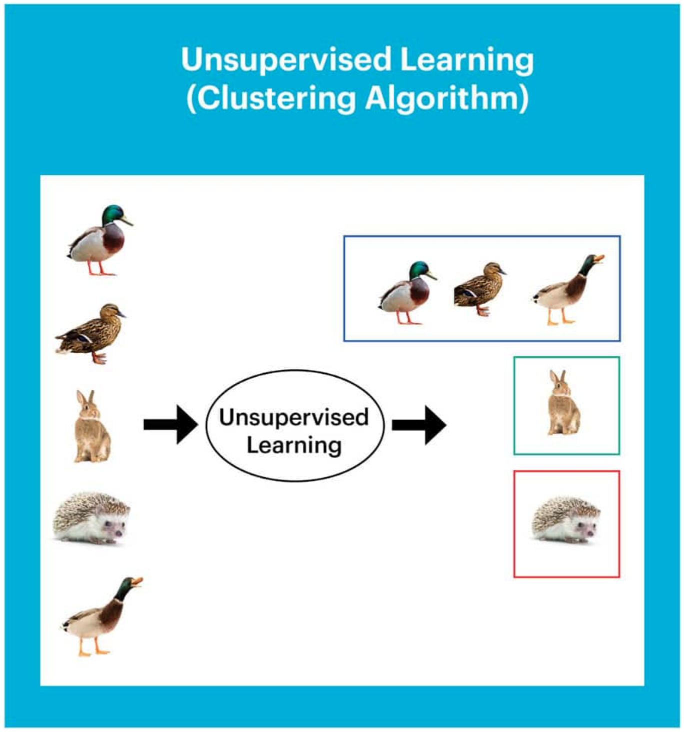

Figure 3, Unsupervised learning enables the machine to identify data with like associations (Zhou, 2018)

Reinforcement Learning

While the supervised and unsupervised learning methods are considered basic, the reinforcement learning method uses no labels and starts with a blank slate. This learning method requires the algorithm to train itself by interacting with the environment (Machine Learning and Automatic Classification, 2018). The algorithm is goal oriented and it learns how to achieve its goal over the course of many steps. It is also incentivized when it predicts information or data accurately. As noted in Domínguez-Estévez's article, "[Reinforcement Learning] algorithms, on the other hand, learn policies that associate "good" actions to execute in different game states based on rewards obtained during the game simulation." For instance, its goal may be to complete a Ms. Pac-Man maze by chomping all of the dots while avoiding the ghosts. The incentive is also to score the highest points. Alternatively, the algorithm can be penalized for being caught by the ghosts, which means the game will end before the maze is completed or the high score is reached--this is the reinforcement aspect. An algorithm that uses the reinforcement learning method may give the appearance of thinking ahead like a human, or like Ms. Pac-Man changing direction to avoid a ghost, which represents the algorithm learning (Domínguez-Estévez *et al*, 2017).

Cultural and Societal Considerations

Principles

The Executive Order on Maintaining American Leadership on Artificial Intelligence (EO 13859) calls for Federal agencies to focus on promoting research of the trustworthiness of AI systems. Section 1(d) states, “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.” The Defense Innovation Board (DIB) is a federal advisory committee for the Secretary of Defense. The Board issued a report that provides principles and an ethical framework for AI systems. The principles state the use of AI systems must be responsible, equitable, traceable, reliable, and governable (Defense Innovation Board, 2019). The trust people have in ML and AI is directly affected by their understanding of how algorithms make decisions.

Computer programmers typically do not develop algorithms with the goal of journaling the predecisional analysis before arriving at the final record, output, or AI decision. In a business process use case, each activity in the course of accomplishing a business process is recorded in an standard operating procedure (synonymous to an algorithm’s if/then statements); however, the actual activity is not typically preserved. Only the final decision or record is preserved after the business process is complete.

Biases and Ethics

AI systems may be subject to biases such as gender, social, cultural, or other ethical considerations. Developers must remain vigilant by ensuring the input data algorithms used are accurate and not skewed, corrupt, limited, or inadequately defined. The inadvertent introduction of human biases and cultural assumptions may result in inaccurate predictive information that may have significant effects.

While there are potential benefits with artificial intelligence, AI use cases should recognize and evaluate biases and ethical concerns. Understanding AI learning methods and how systems are taught are key to identifying inherent human biases and cultural assumptions. AI studies assume gender is binary and language processors often reflect general, racial, and cultural biases. The AI field, similar to Science, Technology, Engineering, and Mathematics (STEM) disciplines, is not diverse. Generally, the information technology professionals are male and either white or Asian (West *et al*, 2019).

The AI Now Institute researches and examines the social implications of artificial intelligence and researches these implications for the following domains: rights and liberties; labor and automation; bias and inclusion; and safety and critical infrastructure. These implications impact the diversity within the AI profession.

AI is currently a technical field, but should be expanded to include other disciplines:

> *AI researchers and developers are engaged in building technologies that have significant implications for diverse populations in broad fields like law, sociology, and medicine. Yet much of this development happens far removed from the experience and expertise of these groups. This has led to a call to expand the disciplinary makeup of those engaged in AI design, development, and critique, beyond purely technical expertise* (Whittaker *et al*, 2018).

In addition to expanding disciplinary representation, the field needs to be demographically diverse.

> *For example, women comprise 15% of AI research staff at Facebook and just 10% at Google. It's not much better in academia, with recent studies showing only 18% of authors at leading AI conferences are women, and more than 80% of AI professors are male. For black workers, the picture is worse. For example, only 2.5% of Google's workforce is black, while Facebook and Microsoft are each at 4%.* (West et al, 2019).

Microsoft's Tay chatbot is an example of an algorithm with racial bias. It was released on Twitter in 2016 (Barbaschow 2019). Tay was supposed to learn from interacting with humans. Within 24 hours, Tay had 50,000 followers and produced approximately 100,000 tweets. Unfortunately, the language Tay acquired was racist and offensive. Tay learned bias from interacting with humans who tweeted offensive language. Amazon Rekognition, launched in 2016, is used by Federal, state, and local governments for facial analysis. The software has been controversial and studies have shown high error rate in identifying non-white faces. Furthermore, the software recognizes only two genders and is unable to categorize nonbinary genders.

COMPAS is an example of an algorithm with social bias, which is used to predict who would be most likely to recommit a crime. The software is a proprietary sentencing algorithm used by judges to sentence criminal offenders. COMPAS tended to over predict that African-American defendants and underpredicted on white defendants being likely to recommit a crime (Broad, *Made by Humans*, 2018).

Ethical considerations must be part of the development process; however, in many cases, this is not being done. Some of the ethical lapses have been caused by the misapplication of data used for training algorithms because there are no standardized methods for documenting datasets. Generally, developers are not experts in machine learning and obtain training data through open access portals. The authors of “Datasheets for Datasets” propose that dataset creators provide datasheets to document datasets. Datasheets are used in industry to summarize performance and other characteristics of a product such a software application or machine component. Similarly, for datasets, the datasheets would summarize the collection motivation and process, composition, and recommended uses (Gebru, 2020).

Records Management Implications

The volume of data created across the globe continues to increase year after year with an estimate of 175 zettabytes being created in 2025 (Stastica 2019). A zettabyte contains 1 million petabytes. As new ways of creating and capturing data are created, the management of the data often lags behind. While the volume can be staggering, the data has to be managed by agencies within a records management framework. The length of retention continues to be driven by business needs and legal requirements. Oftentimes, large collections of data containing personally identifiable information and other sensitive information are brought together for analysis and use. Maintaining the provenance of data through metadata and other methods will be critical for ensuring that privacy and security requirements are met.

Correlation Between Records and Data

The Records and Data Shared Governance Framework, Figure 4, displays Federal records and data as separate but related entities because they are both forms of information. The management of data is overseen by data scientists, systems and software engineers, while records managers are typically personnel with archival, library science, and program management backgrounds. Another difference is typically the fields are disproportionately resourced within each agency and are often not aligned within the same organizational structure. These distinctions mask the fact that the fields are subject to the same governance structures and information retention requirements.

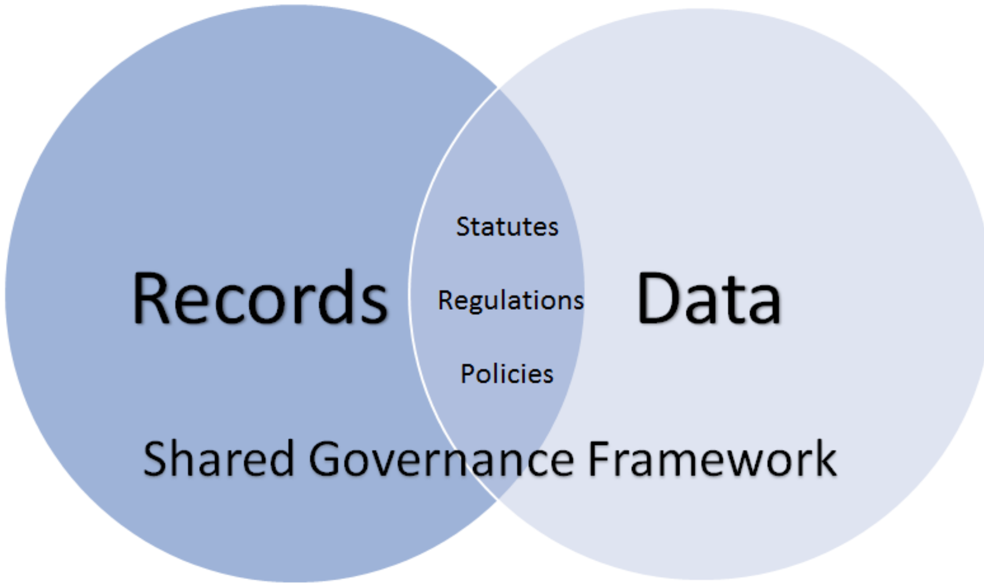

Figure 4. Records and Data Shared Governance Framework

Statutes, Regulations, Policies, and Data Standards

Statutes and Regulations

Records management governance for data is codified in 44 U.S.C. Chapter 33, Section 3301, which states that federal records include "*all recorded information*" regardless of form or characteristics. The term "data", as defined by 44 U.S.C Chapter 35, Section 3502, means *recorded information*, regardless of form. Based on the definition of "recorded information", both records and data share a framework of statutes and regulations.

Policies

The predictive information or output data generated by cognitive technologies, if it is created and documents the business transactions or decisions of an Executive Branch agency, is subject to the Federal Records Act. The transactional data or hashes created on blockchains is an example of how technology may produce records or output data in a non-traditional way (Bhatia *et al*, 2020). Any predictive information or data produced by cognitive technology may be a federal record and agencies are required to manage them throughout their lifecycle.

A key part of NARA's mission is the preservation of historical records. To perform this function effectively, agencies must disposition electronic records at the end of their lifecycle and periodically transfer permanent electronic records to NARA. It is unknown at this point if these cognitive technologies will begin to produce records in non-traditional ways, such as blockchain hashed transactional data.

NARA issues General Records Schedules (GRS), which are equivalent to government-wide records management policy authorizing the management and disposition of federal records. GRS 3.1 for Technology explicitly identifies records related to the general management of technology. Its specificity negates its applicability to cognitive technologies because the algorithms can be designed to produce all sorts of records types and data. Algorithms are not covered by the GRS and should be scheduled by an agency specific disposition authority.

Data Standards

The National Institute of Standards and Technology (NIST) developed a Federal engagement plan for developing standards specific to AI. These standards establish the "rules of the road" for developers when teaching and testing algorithms, which will in-turn drive innovation and new AI capabilities. The plan identifies a number of existing and needed standards, to include non-technical standards to inform policy decisions, such as societal and ethical, governance, and policy decisions. The plan also identifies data standards as a complementary tool to AI standards stating, "Data standards make the training data needed for machine learning applications more visible and more usable to all authorized users (NIST, 2019)."

The Federal Data Strategy (FDS) was established as a result of the Administration's Cross-Agency Priority Goal: Leverage Data as a Strategic Asset. As part of that effort, an intra-agency working group developed a government-wide action plan to identify and prioritize steps to better leverage data to create efficiencies. The FDS Action Plan includes a requirement for the Office of Management and Budget (OMB) to enact a government-wide Data Council responsible for coordinating across agencies on data policy and standards development. Data standards developed by this council and NIST may need to be codified and incorporated into records management regulations, policies, or GRS items.

Authenticity and Integrity

Agency Information Systems Security Officers and Chief Information Officers already engage in risk management techniques to secure information technology infrastructures. Classic controls for managing any risk is to identify vulnerabilities, analyze and evaluate the risk, treat or mitigate, and monitor. Today's technological climate introduces a number of risks to the records and data that cognitive technologies use, including the manipulation of records authenticity and integrity.

Dawn Song, AI security expert and professor at UC Berkeley-

> *...new techniques for probing and manipulating machine-learning systems—known in the field as "adversarial machine learning" methods—could cause big problems for anyone looking to harness the power of AI in business." Song further states, "Adversarial machine learning involves experimentally feeding input into an algorithm to reveal the information it has been trained on, or distorting input in a way that causes the system to misbehave* (Knight, 2019).

The outputs from manipulated algorithms is a breach in the authenticity and integrity barrier for federal records. The outputs at the point of creation have been compromised and thus could adversely affect governmental operations, as well as transactions conducted with the public, and finally the records or data.

Records Appraisal, Scheduling, and Transfer

The emergence of cognitive technologies has not yet precipitated regulatory or policy changes in federal requirements for scheduling and transferring records to NARA. Now that these technologies have become more common within the Federal Government, there are some things NARA and records professionals should consider:

- Records officers and appraisal archivists should work with Chief Data Officers (CDO) for appraisal and scheduling of algorithms and the resulting data sets. The ongoing reuse of data substantiates the need for appraisal and scheduling, as opposed to maintaining data indefinitely. Records retention categories should be based on approved records schedules. Where the application of the appropriate records control schedules to electronic data fields results in multiple options, the CDO, records officer, and business units should coordinate and make a joint decision based on the data characteristics and use. CDOs can lend their technical expertise to determine the business value and retention.
- The fluidity of data makes it difficult to apply records retention requirements; however, scheduling the systems containing the data is more achievable.
- Leveraging AI and ML to identify records eligible for disposition and automating their destruction or transfer into NARA's Electronic Records Archives (ERA).
- NARA may need to review Appendix 2 of NARA Directive 1441, Appraisal Policy of the National Archives, in light of data collected by these technologies and its impact on appraisal.

Conclusion

This white paper provided basic descriptions of AI, ML, RPA, and IoT referred to as cognitive technologies. Various applications were analyzed, to include some enabling factors that support their ability to function. Equally important, the paper provides perspectives on how cognitive technologies may contribute to the automation of records management functions.

Cognitive technologies are driving the exponential growth in the creation and collection of data. It is also important to recognize the learning methodologies enabling algorithms to make human-like decisions or outputs rely heavily on data. AI researchers and developers should implement a framework to prevent the inadvertent introduction of human biases and cultural assumptions that may result in inaccurate predictive information and unintended effects. Federal decision makers who use cognitive technology-produced data should be aware of the ways the data was created, so that policy decisions are informed and documented.

The goal of the records management analysis section is to show that cognitive technologies can be governed within the records management framework. These technologies may impact existing policies and agency standards, such as ensuring records management controls requirements for electronic information systems, and ensuring they adequately maintain the authenticity and integrity of records. Although there have been no records management regulatory or policy changes thus far, this paper highlights the need for forethought on how to impose or tailor policies and standards for data and records created by cognitive technologies.

Glossary

Algorithm	A step-by-step procedure for solving a problem or accomplishing a task.
Artificial Intelligence	The computer science field that deals with the simulation of intelligent behavior in computers or the capacity of a machine to imitate intelligent human behavior.
Cognitive Technologies	Artificial Intelligence (AI), Machine Learning (ML), Robotic Process Automation (RPA), Internet of Things (IoT) or other emerging technology having an impact on federal records management.
Data Cluster	Groups of visual or textual data with similar attributes.
Internet of Things	A device that has a microprocessor and can communicate wirelessly.
Human-computer Interaction	Study of the design and use of computer technology, focused on the interfaces between people and computers.
Machine Learning	Software programming that uses algorithms to autonomously improve decisions with experience or by learning without being explicitly programmed through user interface.
Natural Language Processing	A subfield of computer science for programming computers to process and analyze human language rather than data. Examples include voice recognition personal assistants.
Robotics Process Automation	A technology platform that enables a software robot to interact with applications.
Sensors	Device connected to the internet that collects and transmits data for the surrounding environment.
Software Robot or Bot	An automated program that interacts with system users.

Bibliography

2019 Manufacturing Trends Report, Microsoft Dynamics 365, 2018, available at: https://info.microsoft.com/rs/157-GQE-382/images/EN-US-CNTNT-Report-2019-Manufacturing-Trends.pdf, accessed February 14, 2020.

ACT-IAC. Artificial Intelligence/Machine Learning Primer: Enabling AI and ML Innovation in the US Federal Government, American Council for Technology-Industry Advisory Council, March 12, 2019, available at: https://www.actiac.org/system/files/Artificial%20Intelligence%20Machine%20Learning%20Primer.pdf, accessed March 18, 2019.

Bhatia, Sharmila, E.Kyle Douglas, and Markus Most. "Blockchain and Records Management: Disruptive Force or New Approach?", Records Management Journal, 2020, Vol. 30, Issue 3, DOI: 10.1108/RMJ-08-2019-0040.

Broad, Ellen. *Made by Humans: The AI Condition.* Victoria, Australia: Melbourne University Press, 2018.

Broad, Ellen. Recordkeeping Roundcasts Episode 4: AI, accountability and archives. Recordkeeping Roundtable, October 6, 2018, available at: https://rkroundtable.org/2018/10/06/recordkeeping-roundcasts-episode-4-ai-accountability-and-archives/

Christman, Roger and Kathleen Jordan. Case Study 5: Real World Examples of Email Access Issues, Virginia Governor Tim Kaine Email Project. COSA-NHPRC Symposium, September 15, 2017, available at: https://www.archives.gov/files/nhprc/projects/electronic-records/pdf/case-study-5-real-world-examples-of-email-access-issues-va-jordan-and-christman.pdf

Defense Innovation Board. "AI Principles: Recommendations on the Ethical Use of Artificial Intelligence by the Department of Defense", October 31, 2019, available at https://media.defense.gov/2019/Oct/31/2002204459/-1/-1/0/DIB_AI_PRINCIPLES_SUPPORTING_DOCUMENT.PDF, accessed November 1, 2019.

Domínguez-Estévez, Fernando, Antonio A. Sánchez-Ruiz, and Pedro Pablo Gómez-Martín. "Training Pac-Man bots using Reinforcement Learning and Case-based Reasoning." Paper presented at CoSECiVi 2017, Proceedings of the 4th Congreso de la Sociedad Española para las

Ciencias del Videojuego, Barcelona, Spain, June 30, 2017, Vol. 1957, pp. 144-156, available at: http://ceur-ws.org/Vol-1957/CoSeCiVi17_paper_14.pdf, accessed April 10, 2020.

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. "Datasheets for Datasets", January 15, 2020, available at: https://arxiv.org/pdf/1803.09010.pdf, accessed March 10, 2020.

Hindtze, Arend. "Understanding the Four Types of Artificial Intelligence", *Government Technology*, November 4, 2016, available at: https://www.govtech.com/computing/Understanding-the-Four-Types-of-Artificial-Intelligence.html, accessed March 27, 2020.

"Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025 (in billions)", Statista Research Department, November 27, 2016, available at: https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/, accessed April 10, 2020.

Knight, Will. "How malevolent machine learning could derail AI", MIT Technology Review, March 25, 2019, available at: https://www.technologyreview.com/s/613170/emtech-digital-dawn-song-adversarial-machine-learning/, accessed June 13, 2020.

Lappin, James. "Records management before and after the AI revolution." Thinking Records. January 30, 2020, available at: https://thinkingrecords.co.uk/2020/01/30/records-management-before-and-after-the-ai-revolution/, accessed June 9, 2020.

Lhuer, Xavier. The next acronym you need to know about: RPA (robotic process automation), McKinsey Digital, December 2016, available at: https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/the-next-acronym-you-need-to-know-about-rpa, accessed April 3, 2020.

Miller, Paul. "What is edge computing?", The Verge, Circuit Breaker (blog), May 7, 2018, available at: https://www.theverge.com/circuitbreaker/2018/5/7/17327584/edge-computing-cloud-google-microsoft-apple-amazon, accessed April 10, 2020.

Potter, Matt. “San Diego’s street lights that spy”, San Diego Reader, February 20, 2019, availableat: https://www.sandiegoreader.com/news/2019/feb/20/san-diegos-street-lights-spy/#, accessed January 29, 2020.

Ranger, Steve. “What is the IoT? Everything you need to know about the Internet of Things right now”, *ZDnet*, February 3, 2020, available at: https://www.zdnet.com/article/what-is-the-internet-of-things-everything-you-need-to-know-about-the-iot-right-now/, accessed March 27, 2020.

Reinsel, David, John Gantz, John Rydning. “Data Age 2025: the Digitization of the World, from Edge to Core”, IDC White Paper, US44413318, Seagate, November 2018, available at: https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf, accessed April 10, 2020.

Rolan, Gregory, Glen Humphries, Lisa Jeffrey, Evanthia Samaras, Tatiana Antsoupova & Katharine Stuart. “More human than human? Artificial intelligence in the archive”, *Archives and Manuscripts*, 47:2, 179-203, (2019) DOI: 10.1080/01576895.2018.1502088

“The Rolls-Royce IntelligentEngine – Driven by data”, Rolls-Royce, February 6, 2018, available at: https://www.rolls-royce.com/media/press-releases/2018/06-02-2018-rr-intelligentengine-driven-by-data.aspx, accessed April 10, 2020.

Schuettpelz E, Frandsen P, Dikow R, Brown A, Orli S, Peters M, Metallo A, Funk V, Dorr L. “Applications of deep convolutional neural networks to digitized natural history collections.” Biodiversity Data Journal 5: e21139, 2017. https://doi.org/10.3897/BDJ.5.e21139

Schatsky, David, Craig Muraskin, Kaushik Iyengar. “Robotic process automation: a path to the cognitive enterprise”, Deloitte University Press, 2016, available at: https://www2.deloitte.com/content/dam/insights/us/articles/3451_Signals_Robotic-process-automation/DUP_Signals_Robotic-process-automation.pdf, accessed March 31, 2020.

“ShakeAlert” Earthquake Early Warning System Goes West Coast Wide, U.S. Geological Survey, April 6, 2017, available at: https://www.usgs.gov/news/shakealert-earthquake-early-warning-system-goes-west-coast-wide, accessed April 10, 2020.

Sibley, Graham. Artificial Intelligence and How it Applies to Records Management, Collabware, July 10, 2018, available at: https://www.youtube.com/watch?v=VwnivtEYDdQ, accessed April 10, 2020.

Smith, Ryan P. "How Artificial Intelligence Could Revolutionize Archival Museum Research: A New Study Shows Off a Computer Program's Specimen-Sorting Prowess", Smithsonian Magazine, November 3, 2017, available at: https://www.smithsonianmag.com/smithsonian-institution/how-artificial-intelligence-could-revolutionize-museum-research-180967065/, accessed June 19, 2020.

Thompson, Wayne, Hui Li, Alison Bolen. "Artificial intelligence, machine learning, deep learning and beyond: Understanding AI technologies and how they lead to smart applications" SAS, available at: https://www.sas.com/en_us/insights/articles/big-data/artificial-intelligence-machine-learning-deep-learning-and-beyond.html, accessed on June 19, 2020.

van Essen, Mette. "Machine Learning and Automatic Classification", The Nationaal Archief, December 2019, available at: https://kia.pleio.nl/groups/view/53406652/kennisplatform-innovatie/blog/view/55809165/machine-learning-en-automatische-classificatie-evaluatierapport, accessed February 27, 2019.

West, Sarah Myers, Meredith Whittaker, Kate Crawford. "Discriminating Systems: Gender, Race, and Power in AI." AI Now Institute, April 2019, available at: https://ainowinstitute.org/discriminatingsystems.pdf, accessed April 17, 2019.

Whittaker, Meredith, Kate Crawford, Roel Dobbe, Genevieve Fried, Elizabeth Kaziunas, Varoon Mathur, Sarah Myers West, Rashida Richardson, Jason Schultz, Oscar Schwartz. AI Now Report 2018, AI Now Institute, December 2018, available at: https://ainowinstitute.org/AI_Now_2018_Report.pdf, accessed February 26, 2019.

Wright, Pamela. "Tag it! Introducing the National Archives Facebook Chatbot", NARAtions (blog), October 17, 2018, available at: narations.blogs.archives.gov/2018/10/17/tag-it-introducing-the-national-archives-facebook-chatbot/, accessed April 10, 2020.

Zhou, Linda. "Simplify Machine Learning Pipeline Analysis with Object Storage." Western Digital Blog, May 3, 2018, available at: https://blog.westerndigital.com/machine-learning-pipeline-object-storage/, accessed September, 11, 2020.

Resources

AI Now Institute at New York University, available at: https://ainowinstitute.org/

Executive Order 13859 of February 11, 2019. Maintaining American Leadership in Artificial Intelligence. 84 FR 3867, pp. 3967-3972, Document Number 2019-02544. https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence, accessed March 6, 2020.

General Services Administration. Shared Services, available at: https://ussm.gsa.gov/

M-19-21 Transition to Electronic Records. Memorandum for Heads of Executive Departments and Agencies. Office of Management and Budget and National Archives and Records Administration, June 28, 2019, available at: https://www.whitehouse.gov/wp-content/uploads/2019/06/M-19-21.pdf, accessed April 14, 2020.

National Archives and Records Administration. General Records Schedule 3.1, General Technology Management Records, Transmittal No. 27, January 2017, available at: https://www.archives.gov/files/records-mgmt/grs/grs03-1.pdf

National Archives and Records Administration. Federal Electronic Modernization Initiative, available at: https://www.archives.gov/records-mgmt/policy/fermi

National Institute for Standards and Technology. "U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools", August 9, 2019, available at: https://www.nist.gov/system/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf

Office of Personnel Management. "Reskilling Toolkit: Accelerating the Gears of Transformation", May 22, 2019, available at: https://www.opm.gov/policy-data-oversight/workforce-restructuring/reshaping/accelerating-the-gears-of-transformation/reskilling-toolkit.pdf