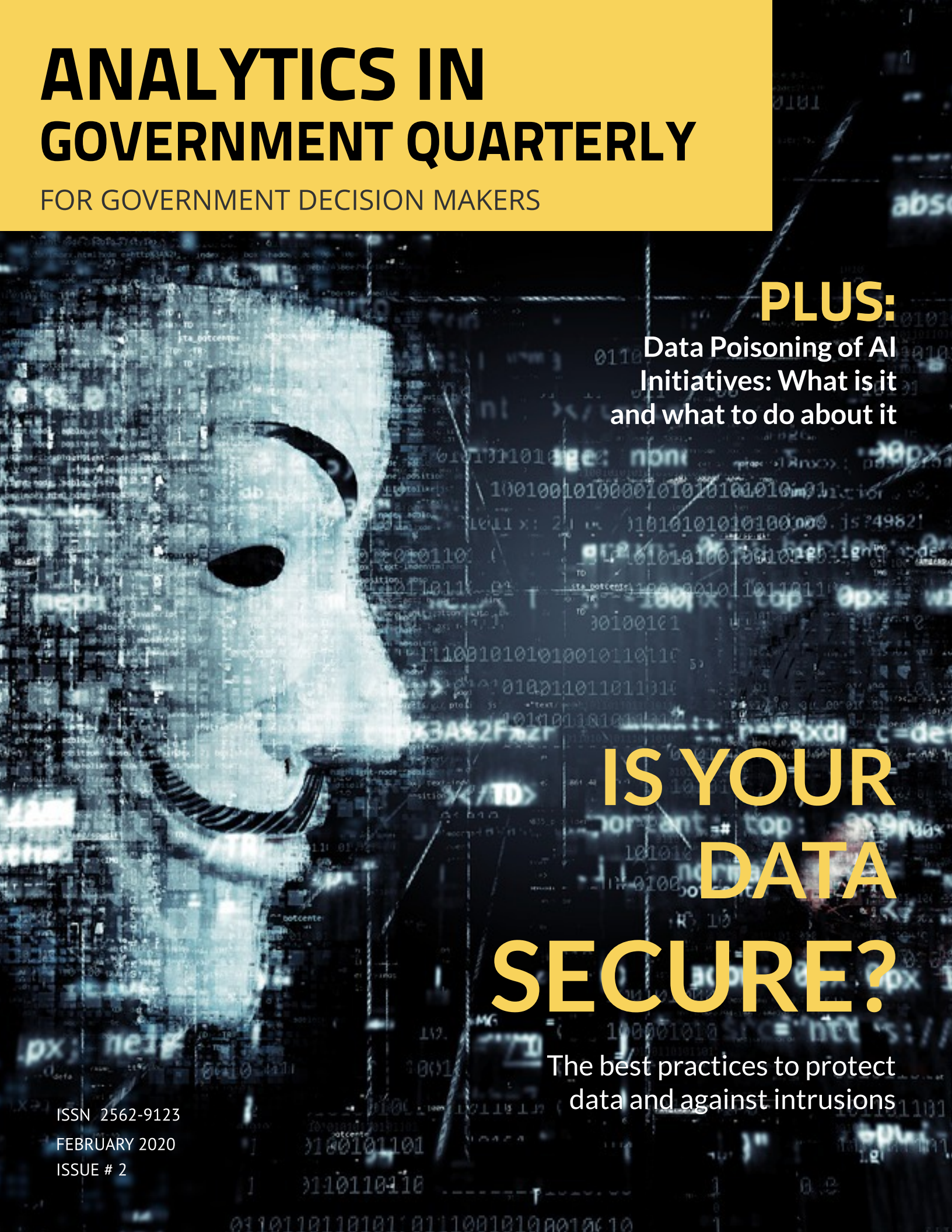
ANALYTICS IN GOVERNMENT QUARTERLY

FOR GOVERNMENT DECISION MAKERS



PLUS:

Data Poisoning of AI Initiatives: What is it and what to do about it

IS YOUR DATA SECURE?

The best practices to protect data and against intrusions

ISSN 2562-9123
FEBRUARY 2020
ISSUE # 2

CONTENTS

A person is looking at a camera.

FEATURES

COLUMNS

ANALYTICS IN GOVERNMENT QUARTERLY

Editorial

Managing Editor |
Dr. Gregory Richards

Editor & Design |
Yvonne Leng

Columnists |
Boris Bogatirev
Wassim El-Kass
Kevin Kells
Hubert Laferrière

Editorial Advisory Board

David Buchanan

Team Lead , Canadian Federal Government, SAS Canada

Stéphane Gagnon, Ph.D.

Associate Professor, Business Technology Management, Université du Québec en Outaouais

Tara Holland

Senior Manager, Government Customer Advisory, SAS Canada

Alex Ramirez, Ph.D.

Associate Professor in Information Systems, Sprott School of Business, Carleton University

Eduardo Rodriguez, Ph.D., MBA, MSC. Math.

Principal IQAnalytics Inc.

Sentry Insurance Endowed Chair in Business Analytics, University of Wisconsin-Stevens Point

Analytics Adjunct Professor, Telfer School of Management, University of Ottawa

Betty Ann M. Turpin, Ph.D

President of Turpin Consultants Inc.

EDITOR'S LETTER

In this Analytics in Government Quarterly (AGQ), we explore issues related to data security. Arnold Toporowski discusses the use of data that include personal information and anonymization approaches that might help relieve privacy issues. Hubert Laferrière examines cybersecurity and the potential of using AI in the battle against intrusions into an organization's network. He points out that research on AI attacks and other anomalies is in its infancy. Kevin Kells explores the importance of human oversight of machine learning algorithms and Betty Ann Turpin discusses a data governance framework that can adjust to continually changing contexts. I contribute an article that examines issues of data poisoning of AI algorithms and discusses the emerging concept of certified external data sets and a "blockchain for AI" idea that secures both the data and the algorithms.

There is no doubt that, as more government organizations adopt AI and Machine Learning approaches, there will be more attempts at intrusions. In the past, these intrusions might be seen as little more than a nuisance that cost time and effort to recover our data or to protect our systems. With AI-enabled decision making however, intrusions could lead to errors in classification of individuals (consider for example, applying for licenses of various types, loans, etc.). Moreover, as we move towards an autonomous vehicle and personalized medicine future, AI algorithms that mistake a stop sign for a speed limit sign or a malignant tumour as benign could lead to dire consequences.

As always, we welcome your comments and suggestions on the articles in this issue. Please do contact us here: agq@governmentanalytics.institute

Gregory Richards

Managing Editor

Corporate

Analytics in Government Quarterly magazine is published 4 times per year by the Government Analytics Research Institute, a consortium of the University of Ottawa, Carleton University, the University of Quebec en Outaouais, SAS and the Institute on Governance. The institute conducts research with government organizations who are experimenting with the introduction of analytics of all forms. Professors and students work on proof of concepts, testing of algorithms as well as examining the organizational practices needed to fully integrate analytics into business processes.

All opinions expressed herein are those of the contributors and do not necessarily reflect the views of the publisher or any person or organization associated with the magazine.

General Inquires

Letters, submissions, comments and suggested topics are welcome, and should be sent to agq@govenmentanalytics.institute or visit our website http://governmentanalytics.institut

Subscription information:

You can subscribe to the magazine online at http://governmentanalytics.institute/magazine/

Reprint Information

ISSN 2562-9123

Balancing Thirst for Data And Privacy Concerns

We expect seamless digital services and online experience, but also expect our personal data to be protected. Is this a realistic expectation in 2020?

Over the last decade, data became a new currency for organizations trying to get ahead and gain a competitive advantage. Data is the driving force behind industrial behemoths like Amazon, Netflix, Tesla, and others. However, with these recent developments, how do organizations balance the need for data to become more profitable and the right to privacy? In that context, multiple questions come to mind to which there is yet a distinctive answer despite multiple and growing attempts by regulators (EU's General Data Protection Regulation - GDPR, Canada's Personal Information Protection and Electronic Documents Act - PIPEDA, etc.) to introduce and enforce more rules for personal data breaches. Moreover, as consumers, we now expect seamless digital services and online experience, while at the same time expect our personal data to be protected and often times not used against us in a covert manner to up-sell, cross-sell, or be shared with affiliates. While compelling, the question becomes whether this is a realistic expectation in 2020.

Improving data security or more specifically, private data security, relies mostly on two strategies.

First and foremost would be improving the organization's security posture, such as reducing cyber-attack exposed surfaces, improving employee's literacy on cyber awareness and phishing, and collecting only the minimal required Personal Identifiable Information (PII) from customers. The second pillar to that strategy, and arguably the more complicated one, relies on improving data management techniques. Growing awareness to privacy concerns and loss prevention of customer's private data is relatively recent. PIPEDA in Canada was passed in 2000, and came into effect in 2004. GDPR in the European Union, arguably a more complete, enforceable and coherent framework, was passed in 2016 and came into effect in 2018. GDPR kick started an avalanche of organizations looking to overhaul their data management practices to protect themselves against the hefty fines set by GDPR.

Organizations have been storing and collecting data, including private data, about customer's behaviours long before privacy laws came into effect or gained meaningful enforcement mechanisms. For long periods of

time data (including PII data) was neglected and not protected by many organizations, which in turn, now find it very complicated to untangle that history of neglect. Improving data management requires enhancements to:

- Metadata Management – To distinguish sensitive data from non-sensitive data;
- Data Cataloguing – To properly identify and classify systems and repositories that collect and/or store data;
- Data Governance – To ensure issues are being remediated when found in IT Architecture; and
- Principles – To introduce data masking, row level security, encryption of data repositories.

Before the introduction of GDPR, these were unpopular activities with undefined return-on-investment, high complexity, lack of quick and tangible wins and mostly no apparent business case. Since GDPR introduced heavy fines on a per-row (customer) basis, many companies started to implement risk mitigation strategies to improve their legacy systems data management practices. It is important to remember that data monetization is not the opposite of data security. On the contrary, many leading businesses that introduce modern data security postures such as privacy-by-design, open-by-default-closed-by-exception are leaders in monetizing their data.

Privacy by Design

In this complex electronic business environment, a "check the box" compliance model leads to a false sense of security. That is why a risk-based approach to identifying digital vulnerabilities and closing privacy gaps becomes a necessity. Once you have done the work to proactively ensure that your controls are implemented and your information is secure, having the privacy practices verified against a global privacy standard can take the organization's privacy and security posture to the next level. Privacy by Design means building privacy into the design, operation, and management of a given system, business process, or design specification; it is based on adherence with the 7 Foundational Principles of Privacy by Design:

1. **Proactive not reactive—preventative not remedial.** Anticipate, identify, and prevent invasive events before they happen; this means taking action before the fact, not afterward;
2. **Lead with privacy as the default setting.** Ensure personal data is automatically protected in all IT systems or business practices, with no added

A black CCTV on the wall.

A door with private signage

action required by any individual;

3. **Embed privacy into design.** Privacy measures should not be add-ons, but fully integrated components of the system;
4. **Retain full functionality (positive-sum, not zero-sum).** Employs a "win-win" approach to all legitimate system design goals; that is, both privacy and security are important, and no unnecessary trade-offs need to be made to achieve both;
5. **Ensure end-to-end security**. Data lifecycle security means all data should be securely retained as needed and destroyed when no longer needed;
6. **Maintain visibility and transparency—keep it open.** Assure stakeholders that business practices and technologies are operating according to objectives and subject to independent verification;
7. **Respect user privacy—keep it user-centric.** Keep things user-centric; individual privacy interests must be supported by strong privacy defaults, appropriate notice, and user-friendly options.

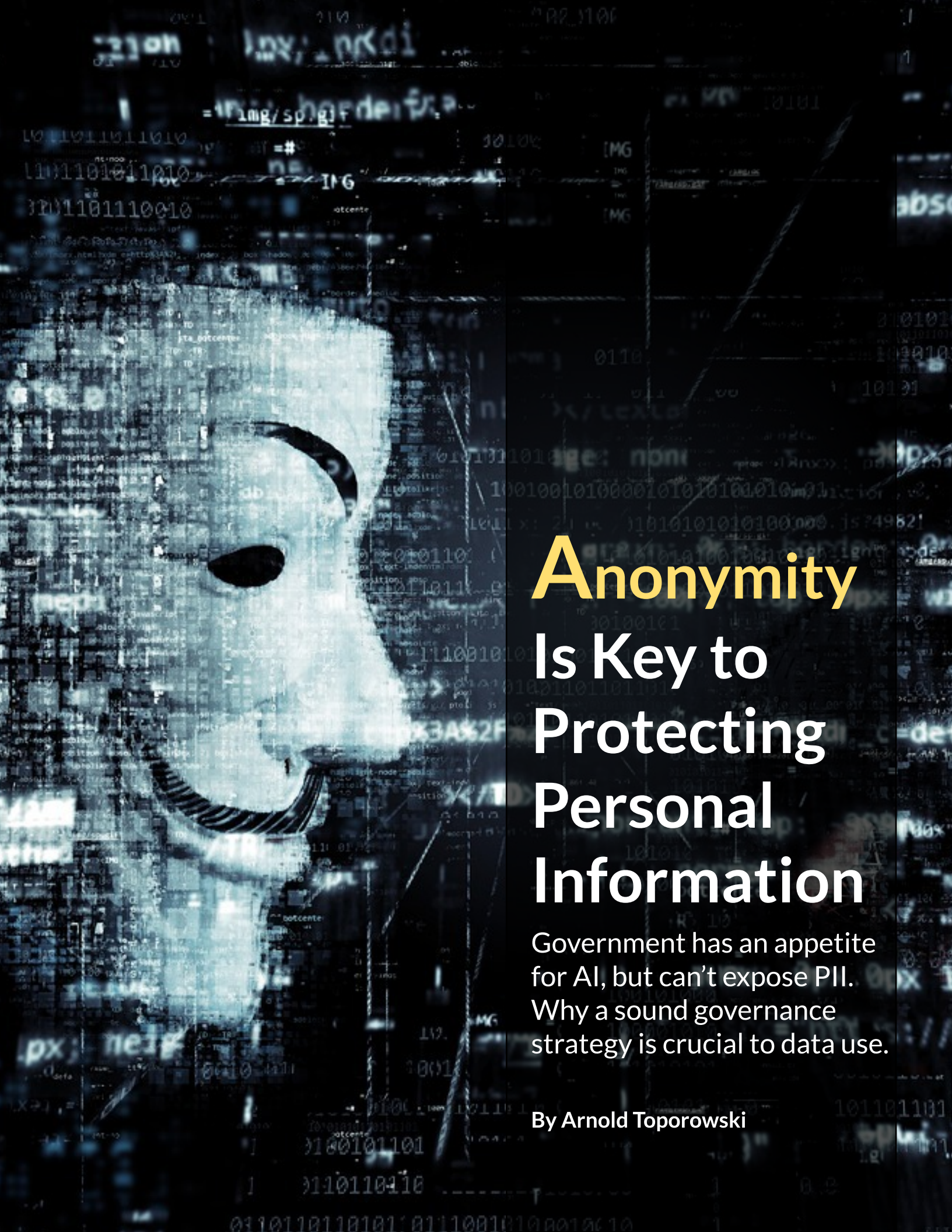
Dr. Ann Cavoukian, Executive Director of the Privacy and Big Data Institute at Ryerson University, Three-term Information and Privacy Commissioner of Ontario is concisely summarizing the topic:

"Protecting privacy while meeting the regulatory requirements for data protection around the world is becoming an increasingly challenging task. Taking a comprehensive, properly implemented risk-based approach where globally defined risks are anticipated and countermeasures are built into systems and operations, by design can be far more effective, and more likely to respond to the broad range of requirements in multiple jurisdictions."

About the Author

Boris Bogatirev is the Ottawa Marketplace Leader for Omnia AI, Deloitte Canada's AI Practice. He is helping Canadian organizations to embrace and leverage AI to leapfrog their competition and to provide better services to all Canadians.

Boris has 20 years of experience with the entire breadth of data applications from strategy to implementations. Boris led multiple AI-centric projects with leading federal departments and



Anonymity Is Key to Protecting Personal Information

Government has an appetite for AI, but can't expose PII. Why a sound governance strategy is crucial to data use.

By Arnold Toporowski

Governments are eager to explore the benefits of artificial intelligence. AI has the potential to streamline services and provide better risk management for government departments. But there's also the potential to expose personally identifiable information (PII)—a violation of federal privacy legislation.

For example, the delivery of benefits like employment insurance and disability payments is a labor-intensive process. Applications and claims must be screened for eligibility (and possibly misuse). AI processing can speed the assessment of claims and improve the accuracy of service delivery.

Better risk assessment also promises improvements in efficiency and effectiveness. AI has a proven track record in risk assessment for financial institutions, one that could be applied to processes in the Canada Revenue Agency, for example. Zero-adjustment audits—audits that find no discrepancies—are frequent, and a drain on resources. Analytics that can identify higher-risk returns and prioritize them would allow auditors to spend their time more effectively.

Appetite for Data

If government has an appetite for AI, artificial intelligence has an appetite for data. Fortunately, that's a hunger government can feed. With a body of 50-plus years of digitized records and a statistical agency that's the envy of other nations, government is well-equipped to deliver on the data front. While the historical data may be incomplete by today's standards, e.g., with new fields and data formats like video and audio, it's a boon for machine learning (ML). ML is a key piece of the AI puzzle, allowing systems to learn patterns without human intervention.

There are many third parties who are interested in having access to government data and who would be able to provide government with valuable insights and results. Some government departments are very interested in making some data public for data science projects.

With the rise of post-secondary data science programs at Canadian universities, there's an opportunity for government to share databases for post-graduate projects that can have an impact on service delivery.

But there's a catch: The most useful data contains personally identifiable information that can't be disclosed to third parties. The federal government's Personal Information Protection and Electronic Documents Act (PIPEDA), along with similar legislation at the provincial level, governs how organizations can collect, use and disclose personal information. The use of PII is tightly controlled.

One solution is to use aggregate data—data amalgamated from a large number of personal records with the PII omitted; while that may be useful for predicting trends in service usage, there are circumstances under which that's not useful.

Anonymization

If government wants external data scientists to create models that can predict behavior of individuals, aggregate data isn't useful.

Enter anonymization—redaction, encryption or obfuscation of PII that maintains the integrity of the individual record.

It's more complex than it sounds. The Canadian Institute of Health Information has said that that encryption of a patient's health

Personal data - a finger print

card number, name, address and telephone information is inadequate. An individual's records can very likely be identified with just birth date and postal code alone.

This is called low-frequency data. Statistics Canada has a process in place to blank out such data. There's a minimum threshold for the number of results returned by a query. Below that number, results are excluded. While Statistics Canada largely deals with aggregate data, it becomes more granular the further you drill down into a particular cell.

So how do you maintain protection of PII and still get useful results?

One approach is to broaden the scope of particular fields. For predictive power, a person's age may be important, but the actual date of birth is not necessary. The address region may be significant, but the specific postal code might not be. Postal code K2J 1S8 is a stretch of houses on Greenbank Road in Nepean; the K2J forwarding station has 23,593 delivery locations. Using only the first three characters of the postal code makes personal identification much less likely. Anonymization is a sliding scale. This fuzziness of data helps protect PII.

Best Practices for Securing PII

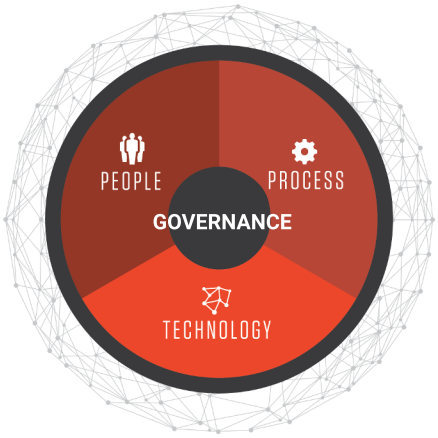
Like anything related to information security, a PII governance strategy relies on three elements: people, processes, and technology.

People. A data governance committee, chaired by a chief data officer or chief information officer, establishes the principles of data use. How can the data be used? What data can be disseminated, and for what purposes? Who can use what data? This strategy must abide by the guiding principles of PIPEDA. Data stewards are responsible for ensuring access is according to the established protocols.

Process. The process for data use must secure critical information, but it must be usable as well. Automated processes should be guided by the Canadian federal government's Directive on Automated Decision-Making, a first-of-its-kind framework for determining when human intervention is required in automated processes. At the same time, it can't be overly bureaucratic—the objective is to streamline freeing up users for more value-added activity.

Technology. The right technology for governance is also critical. The data governance strategy must be embedded in the technology, automatically applying governance rules to access to data. For example, technologies like SAS Data Preparation and SAS Federation Server can govern data access and put in place secure views that prevent access to raw fields as necessary.

Increasingly there are reports of stolen personal data from corporate systems showing that



the exposure of personally identifiable information is a real risk. However, that mustn't stop government from using data to provide better, faster and more effective services to its citizens. With a sound, well-designed governance framework embedded in processes and technology, and the promise of AI and machine-learning technologies to free up staff for more value-added activities than scrutinizing claims and applications, these opportunities can be exploited with controlled risk.

About The Author

Arnold Toporowski, Government Customer Advisory at SAS Canada, brings over 30 years of experience in Information Technology specializing in DataOps to the Government Customer Advisory Team.

Data Poisoning of AI Initiatives: What is it and what to do about it

Data poisoning occurs when an attacker injects bad data into your model's training data set. This approach tends to degrade the overall AI model leading to erroneous results.

By Gregory Richards

Data security is an important area of research that involves new methods of authentication (e.g., multi-factor authentication), secure sites (e.g., blockchains) and cybersecurity solutions (e.g., encryption).When it comes to AI however, less is known about potential attacks, but more research on data poisoning is emerging that is helping data scientists and managers improve security of their AI algorithms. It is an important issue because in some cases, such as autonomous vehicles, medical imaging, or precision medicine, these attacks can be downright dangerous.

This article reviews some of the recent research on the types of attacks being noted in the field and potential solutions which include ensuring proper data validation and cleaning, as is common in most AI projects, but the article also considers new ideas such as certified data sets and the use of cryptography for integrated management of the data pipeline.

Basics of Standard AI Algorithms

To understand how data poisoning works, we need to briefly review how an AI algorithm works. Neural networks are the most frequently-used approach for implementing AI. These networks can become extremely complex but fundamentally, the entire system is a predictive model (or a series of predictions in the case of deep learning). If, for example, we wanted to predict house prices (the response variable) based on square footage (the predictor variable), we'd need thousands or hundreds of thousands of examples of prices and square footage.

The figure below tells this story graphically. The blue line shows the actual relationship of actual and predicted data observed in

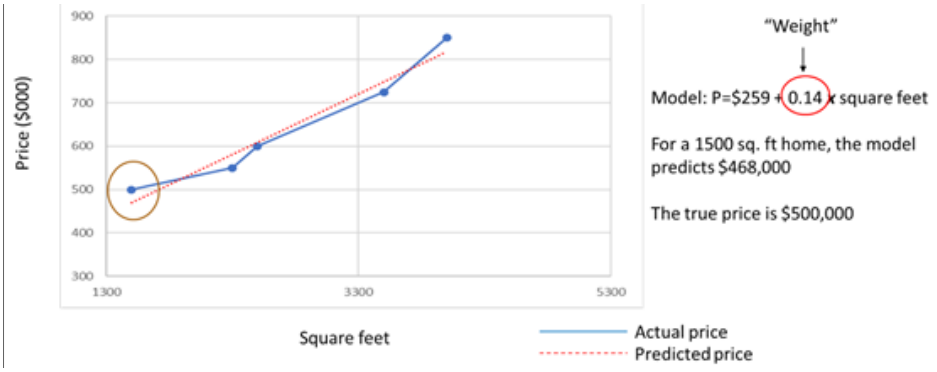


Figure: a model for house prices prediction.

our data set. But we want to predict house prices for square footage of houses that are not in our data set. To do this, we'd like to find a model that reduces prediction error (represented by the distance between the blue and the red lines in the graph above). This model's equation is shown on the right in the figure.

To create the equation, the data scientist would gather a large data set with thousands of examples of square footage and the associated house prices. The scientist will use some of the data to train the model, leading to identification of the intercept and weight that we will use for our prediction. The data scientist would then validate this model by using it to predict the prices of houses in the remainder of the data set. Here we are predicting a quantitative variable (the price of a house), but the same approach applies for classification engines such as image recognition (e.g., classifying animals as cats or dogs, or pictures of tumours as benign or malignant).

What is data poisoning?

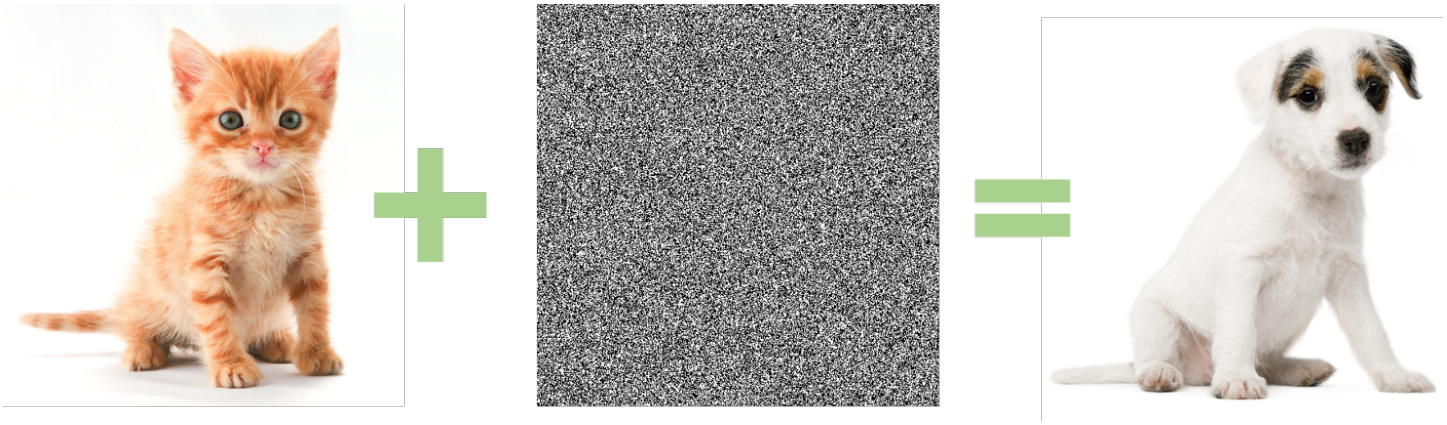
Poisoning occurs when an attacker injects bad data into your model's training data set. This approach tends to degrade the overall AI model leading to erroneous results.

How might this be done? Consider that nowadays, a lot of the data gathered for training AI algorithms are gathered through open sources: social media sites, published data for image recognition etc. Data poisoning can therefore occur if attackers access these sites and inject erroneous data. For training data within an organization, the attacker would need access to your training set, a situation that is less likely (but still possible if someone inside the organization has an axe to grind, or your system is externally hacked).

Adversarial attacks occur when the attacker inserts a new sample into the training set that pollutes the relationship between the predictor and the response variable. For image recognition classification systems, the output can sometimes be very far from the input as shown below when our kittens become misclassified as puppies.

Another approach is referred to as backdoor poisoning attacks which tend to be more targeted and therefore more insidious. In the first case, where the training set might be compromised, the algorithm might converge poorly or not at all. In most instances, the data scientist can usually tell something is not right. In the second case, targeted attacks by contrast, can be difficult to detect. Research in the field of image recognition suggests that even exceedingly small changes can cause the system to misclassify the input variables.

Ilja Moisejevs' "Towards Data Science" blog provides additional examples that includes logic corruption where the attacker changes the algorithm. This could result from someone inside the organization who has access to the system on which the AI algorithms are being



Kittens may be misclassified as puppies when adversarial attacks happened. Image Source: Francesco Gadaleta's blog post

A person is using vacuum to clean the carpet

developed. But, as AI tools become more widely deployed, we can expect that hackers will become more prevalent and more determined, therefore any hacks to the organization's network might also compromise their AI implementations.

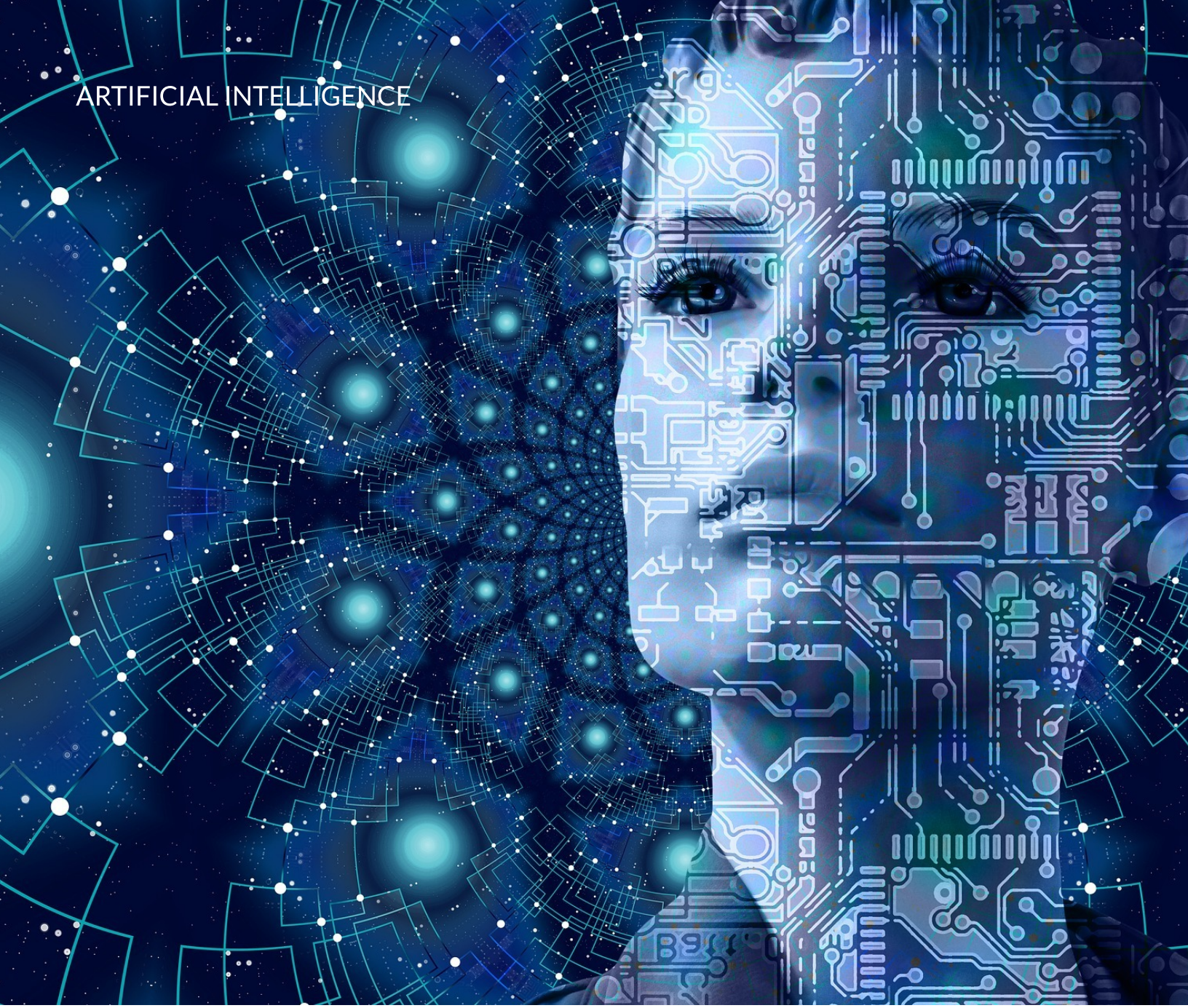
What to do about it?

Data cleaning is a key step in most data science projects. Here the defense is one of paying close attention to outliers or to algorithms that take a long time to produce results. In addition, it is important to ensure that if you are using external data to train algorithms, the data has been secured and cleaned before integration into the training set. The message here is that management of the overall data pipeline, from data definition and formatting, to capture through to implementation, must be a top priority for organizations applying AI, particularly if the tools are being used to automate decisions.

A new idea emerging from some researchers is the notion of a "fitchain". The fitchain idea includes three core elements: certified training data sets, cryptographic proof of the data set and of the training model, and storage of the meta data and model in a blockchain-like public ledger. This approach aligns well with the concept of data pipeline management. In this case, critical components of the pipe include immutable ledger technology that alerts users to any changes to either data or to the algorithms during development of AI solutions. As AI becomes more popular, it's no doubt that attempts to hack the algorithms will increase. An integrated data pipeline approach (not only for AI but for all types of analytics in the organization) would help keep managers abreast of the potential for harm, as well as steps they can take to avoid negative consequences.

About The Author

Gregory Richards, MBA, Ph.D., FCMC, is currently the Executive MBA Director and Adjunct Professor at the University of Ottawa. He was a visiting professor at the Western Management Development Centre in Denver, Colorado and a member of Peter Senge's Society for Organizational Learning based at MIT. His research focuses on the use of analytics to generate usable organizational knowledge.



AI AGAINST AI

The specific strategy to protect AI products and systems against cyber-attacks is in its infancy: data scientists and the Security IT team must put on their creative Seven League Boots to deal with the fast-moving threat.

By Hubert Laferrière

Cybersecurity is of the outmost importance for the AI community. AI is now considered as a vector of crime by many: with the democratization of AI where members of the public have gained access to key resources needed to use and develop their own AI tools (data, software, and hardware) comes the empowerment of malicious actors to use AI for nefarious purposes [1].

From the average computer user to programmers and coders, the key known challenge everyone faced, was, and still is, about cyber-attacks on computers and systems, government systems included. In the summer of 2019, an IBM Study showed that data breach was on the rise and can cost the average business up to $3.92 million[2]. At Immigration, Refugees and Citizenship Canada (IRCC), the specific strategy to protect AI products and systems against cyber-attacks is in its infancy: data scientists and the Security IT team must put on their creative Seven League Boots to deal with the fast-moving threat.

Defense the Integrity

When it comes to AI and cyber security, new challenges or vulnerabilities are emerging. The AI community is now assisting in the confrontation of the machines: AI against AI. From the perspective of the IRCC AI team, this would mean attacks on the integrity of algorithms.

Mr. Jose Fernandez, Associate Professor at Polytechnique Montréal outlined potential outcomes at the Symposium on Algorithmic Government, organized by IRCC in April 2019: "The lack of explanation provided that many Machine Learning-based AI solutions that can lead to unconscious bias, hidden manipulation." [3] So, any attempts to "game" the algorithm that we had developed for instance may generate dire consequences.

The IRCC project started over two years ago (the project is described in the first edition of the AGQ Magazine, November 2019). The AI team at that time concentrated efforts on building a machine learning model that is a solid proof-of concept, ensures the best predictive models and, at the deployment phase, the model performs adequately error-free. The team faced many challenges, namely the ethics, privacy and legal aspects; the team ensured the algorithmic models were void of harms, such as discrimination generated by bias, that they respected procedural fairness principles, and that the bottom line outcomes focused on improving productivity enabling more agile, flexible and fast operational processes.

Gaming the Model

Cybersecurity was not at the forefront of the team efforts, although a systematic monitoring process (algorithm robustness) to detect odd patterns was in place, patterns that led us to determine whether "gaming" activities were happening. When such an event occurred, the team immediately modified parameters, maintained a higher level of monitoring and conducted some tests.

François Nadon, Director of the IT Security and Production Services at IRCC, proposed to hire a student to game the models, more precisely to determine the degree at which one of the models was at risk. The student used different ML algorithms and was able to identify rules that had the potential impact of misclassifying applicants.

"The lack of explanation provided that many Machine Learning-based AI solutions that can lead to unconscious bias, hidden manipulation."

Jose Fernandez, *Associate Professor, Polytechnique Montréal*

1. Dupont, B., Stevens, Y., Westermann, H. & Joyce, M., 2018., Artificial Intelligence in the Context of Crime and Criminal Justice. A report for the Korean Institute of Criminology.
2. IBM. (2019-07-23). IBM Study Shows Data Breach Costs on the Rise
3. Fernandez, J., 2019 April, Artificial Intelligence and Cyber security: Challenges

Although the impact was limited, the team immediately reviewed parameters and rules.

We are now working on developing a method for better assessing gaming risks for every machine learning model, a method that would be systematically integrated into AI processes and procedures. This is a must for ensuring the model's integrity.

AI & Increased Vigilance

In our quest to find solutions to counteract gaming threats, we were looking for existing ones that could be found in the marketplace such as automated model products for monitoring any attacks or consultant expertise to help us learn how artificial intelligence be used to increase cybersecurity. The ideal solution is to have AI to support the work of both cybersecurity analysts and AI teams by detecting an anomaly. To our astonishment, little or no existing models exist and we had difficulty finding qualified experts. According to Mr. Fernandez, only a few research groups in the world are working specifically on securing AI systems[4].

Dave Masson, director of the Canadian division of a British firm specializing in this field, points out that AI can equip analysts and, as a result enable them to react more quickly in the event of a cyber-attack[5]. He is convinced that the contributions of AI will continue to increase because of the ever-increasing number and complexity of cyber threats. As such, it seems that it's almost impossible to keep up with threats and it is increasingly obvious that we will need to use AI to stay in the race.

In the meantime, monitoring and testing activities appear to be the best means at our disposal to increase our vigilance. To do so, the IRCC AI team must mobilize resources for constant monitoring and development, as needed, and identify more sophisticated approaches to improve detections. The team will still hire students to game the models. We resume work with our IT Security colleagues and our quest to find a solution in the marketplace. At the end of the 2019 Symposium on Algorithmic Government, some participants suggested the creation of a dedicated working group of civils servants and academics on AI and cybersecurity. That could be a path to tackle the challenge, at least to open a dialogue among members of the AI community. The IRCC AI team may take the lead assuming there is enough interest in the AI community for moving forward.

About The Author

Hubert Laferrière has established the Advanced Analytics Laboratory for the Department of IRCC. The Lab has just recently been transformed into a centre of excellence for AI under the name of Advanced Analytics Solution Centre (A2SC). He is currently leading a major transformative project where advanced analytics and machine learning are used to augment and automate decision-making for key business processes.

4. Corriveau. É. (2019-05-25). Cybersécuritaire, l'industrie 4.0?
5. Rettino-Parazelli,K. (2018-04-06). L'intelligence artificielle pour accroître la cybersécurité.

By Kevin Kells

Use Data Science Wisely by Incorporating Human Due Diligence

Your organization is poring through a pile of résumés of candidates for an open position, hoping to choose a new hire who will be a good fit, will excel at the work, and will stay with your organization for a longer time. It's hard to know answers to these questions from the résumé content. Is there a way to use data science to prioritize the best candidates and to screen out poor matches?

The obvious answer is "yes." Supply résumés from prior years' candidates to a Machine Learning algorithm. Tie each résumé with pertinent facts about their subsequent career trajectory at your organization, where longevity and achievement are scored higher. Having trained the model, now you're ready to feed the new candidate résumés to the model so that it scores each one.

This article is not about how to do that. It is about the caveats you need to be aware of when applying data science: Validity, Impartiality, and Transparency.

In the book, "Weapons of Math Destruction", author Cathy O'Neil reveals important aspects of data science that when not properly understood and corrected for, can lead—and have led—to unintended, negative consequences. Some individuals have had their lives ruined and communities have been inadvertently discriminated against by the inexpert application of data science.

When setting up a system that depends on data science, build in a process of human due diligence through the following five steps:

First is statistical validity. Let's say we've trained our résumé Machine Learning model and our algorithm has scored the new candidate résumés with a value from 1 to 10. If the number of résumés and career data we've used for training is small, the accuracy may be so fuzzy that there is no meaningful difference between a score of 7 and a score of 10. If so, then an alternative scoring should be adopted, say just three piles of résumés: likely, possible, and unlikely job matches, so that undeserved preference is not ascribed within each pile.

Second is to continually measure the model against reality. If it predicted "likely job match" for an individual last year, how well is that employee doing in their career now? What score would your organization have given the résumé then with the benefit of hindsight now? What score did the model actually give it? The model should be updated and retrained based on continual human audits and reality-checks.

Third, question whether what you are measuring is what you want to measure. If the career data that trained the model includes the length of employment and number of promotions, then that is what the model will tend to score higher


A person is checking his data

in the new-hire résumés. What about the genius employees who in the past made a lasting contribution well beyond expectations though they left sooner than expected to do bigger and better things? A model based only on promotion count and employment length may end up placing genius candidates in the "unlikely match" pile. Review continually what you are telling the model to score.

Fourth, avoid bias by having a disinterested party review what you are trying to achieve with the model; this will permit you to assess blind spots and unintended prejudice. Imagine a crime prediction model which integrates the location where serious crimes and petty crimes are committed onto a map to direct where police patrols should spend more time. Perhaps serious crimes in the city occur more often in certain hotspots unrelated to the residents of the area, while petty crimes, which may go unreported without a police presence to witness them, perhaps occur more frequently in poorer areas. The patrols sent into an area by the model who witness petty crimes will report them to the model. The increased reports, in turn, increase the patrols sent there. There is a feedback loop which may bias the model to send more patrols to poorer areas instead of to serious crime hotspots. Such a propensity to bias should be assessed, and the model adjusted to best serve the highest good of the whole city fairly.

Finally, if your model provides a score, make that score transparent. Your users should be able to click through to see a breakdown of the factors affecting the score, to read the formulas or methods used to calculate the score and its components, and to learn which data was used to train the model and how that data was prepared. Could such disclosure reveal proprietary information or permit individuals to "game" the model? Perhaps, but weigh those downsides to the damage the model may inflict on an individual's life and the number of individuals affected by the model.

Data science should be used wisely, and the incorporation of human audit and review processes is essential to avoiding mistakes, especially when they can affect people's lives. Human due diligence can ensure the validity of a model is understood and properly used within its limitations. The model should have an update process that compares predicted values with actual values on a continual basis to ensure adjustments are made to improve the model. Awareness that unintended bias and hidden feedback loops may exist is important, and the enlistment of a third-party review of the aims of the model and possible hidden biases is a way to ensure impartiality of the model. Finally, providing transparency in how a score is calculated and what methods and data are used plays a vital role in mitigating unintended negative effects on the lives and livelihoods of real individuals that a single, unexplained score might otherwise cause.

About the Author

Kevin Kells has worked as an R&D Engineer in software systems in the Financial and Semiconductor industries in Switzerland, Silicon Valley, and Ottawa, and currently works with real-time data and news feed systems at a major market news and data company in New York City. He also has extensive experience in non-profit management, both in the area of human systems and IT systems. He received his PhD from the Swiss Federal Institute of Technology (ETH), Zurich in computer simulation of semiconductor devices and holds an MBA with areas of focus in entrepreneurship and business analytics from the University of Ottawa, Telfer School of Management.

Key Considerations in Establishing Sustainable Data Governance

Organizations are in a constant state of flux, even with a 3-5 year strategic plan "in-hand". Today, constant changes are driven not only by external factors, but by the greater reliance on information management and technology systems (IM/IT). With data at the heart of many government digital transformation initiatives, sustainable data governance is paramount. This article explores how data governance practices and key lessons can help to address this challenge in a continually changing environment. In strictest terms, governance entails authority, decision-making, and accountability[1]. The definitions of data governance are broader and several exist. Some are simple:

- a collection of practices and processes that can support formal management of data assets;
- others are more comprehensive including data quality, stewardship, security and privacy, usability, integration, compliance, roles, integrity, tools and skills.

Inclusive data governance frameworks

Regardless of the complexity of the data governance framework adopted, a shift is needed from the former top-down, IT focused and workflow-driven[2] approaches to more eclectic and inclusive frameworks such as the SAS best practice model[3], which is a comprehensive approach to data governance sustainability. The attraction of the SAS model is its comprehensiveness as that integrates several aspects to enable sustainability. Figure 1 incorporates the main elements of the SAS approach, as well as a few other considerations (See the next page).

The SAS model is comprised of three primary dimensions:

1. **Community** – an inclusive body of individuals at various levels and functions, whose interest is data. Ensure that all business areas are accountable for and share responsibility for success (or failure). Create transparent awareness and knowledge about data at all levels in the organization that helps to manage expectations.
2. **Environment** – invests in and develops the knowledge, skills, and tools to manage data as an

1. Institute on Governance. (2020). Defining governance
2. Chisholm, M. andKalb, A. (2018). Agile Data Governance:a Bottom-up Approach.
3. Nevala, K. (unknow). Sustainable Data Governance:A SAS Best Practices

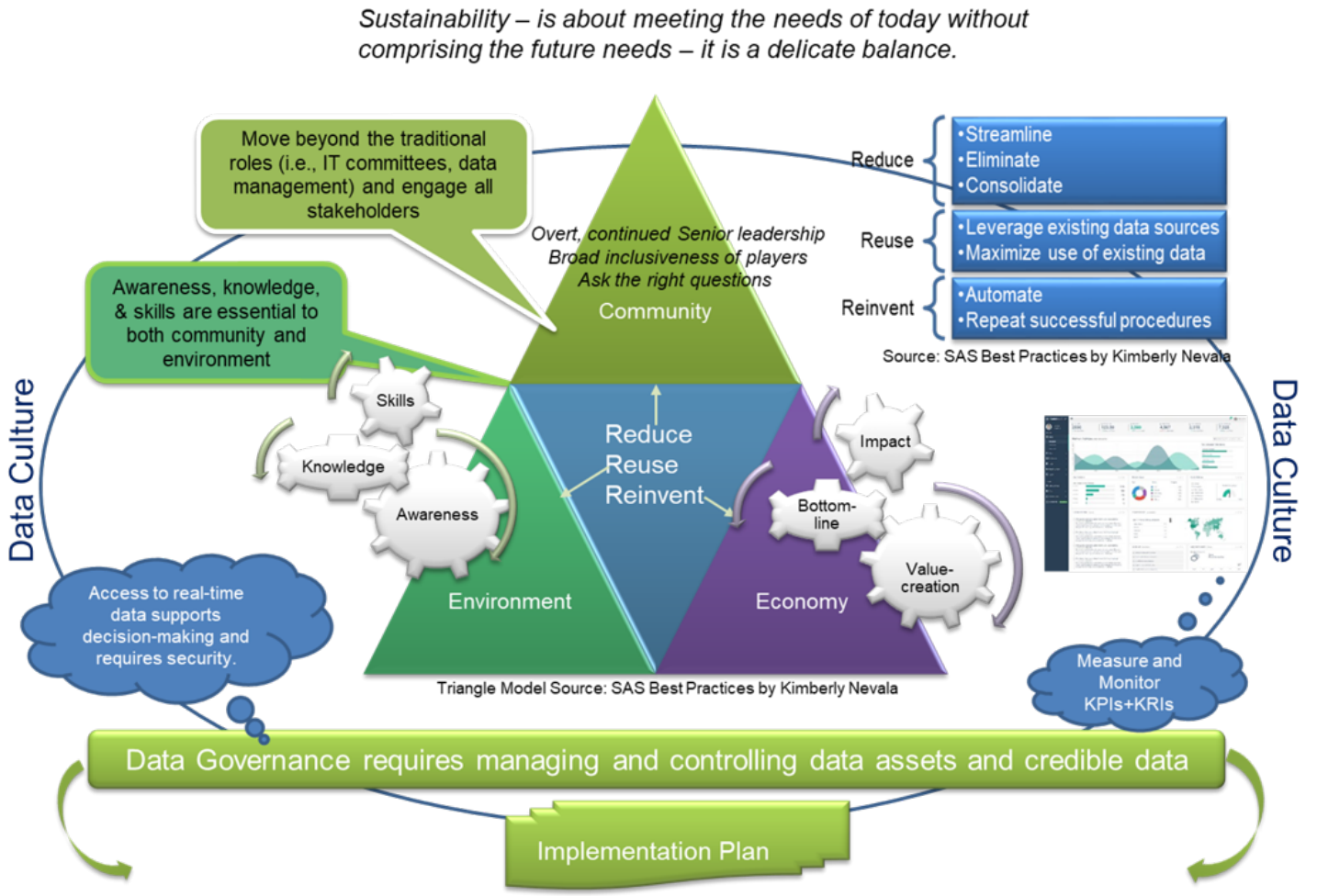


Figure 1: Data Governance Sustainability Universe

asset. Defining roles, responsibilities, and authorities for data. Governance has a long-term role in the strategic decision-making in the organization.

3. **Economy** – is about incorporating value creation, which means relying on data governance as a means to the strategic goals of the organization. Measurement is essential.

The SAS model goes beyond simply addressing IT management and assets. Each dimension calls for defining bodies and approaches for ensuring authority, decision-making and accountability. These SAS primary dimensions have been identified, in general terms, in various government data analytic strategies and/or, in part, in data culture or data literacy plans and renewed governance committees. In some cases, the dimensions are found in separate plans, which raises the concern about fragmentation, despite consultation efforts, because integration of all dimensions within the organization is necessary for sustainability.

How to apply frameworks effectively

The concept of sustainability necessitates that the framework be evergreen to adapt to ongoing changes in the organization, a necessity for sustainability success. A few key lessons have been learned in applying these types of frameworks.

First, within the community dimension, it is essential to have visible, active and ongoing senior management commitment. A major role for senior management is to manage competing priorities and increasing demands from central agencies, which often occur with little to no claw back on previous expectations.The priority focus must shift to data governance as the single most important priority, second only to the people (employees and

citizens). Second, experience has shown that committees/governing bodies create typical terms of reference (i.e., mission, roles, standards, members, model, policies/procedures) but often the goals are too lofty or vague, entailing too many levels of approval to be adaptable and timely; and exclude the functional experts.

Third, many strategies/plans or governing body terms of reference do not follow-up with an implementation plan. This oversight can lead to the plans and/or the data governance body becoming static or relying on a sub-working groups who are not resourced or empowered to effect change. As well, most data governance plans usually only address the IT asset needs. With an implementation plan, the organization will be able to position data governance as a long-term initiative ingrained into the fabric of the organization's business.

Finally, a plan is not worth much if committed resources are not attached. The tendency in government is to do more with less, do it on the 'side of the desk', or if we can get to it. This will not suffice. 'Reuse, reduce, and reinvent" are part of the SAS model. Resource efficiency can also be achieved with streamlining business processes and extraneous decision-making steps[4]. As well, the re-allocation of resources is a logical solution while ensuring that organizational strategic outcomes are not compromised. It is also crucial that management ensure that sufficient long-term resources are available to finance all roles and functions required for robust, ongoing data governance.

What can influence data governance sustainability

The foregoing arguments suggest that the data governance committee must explicitly define the right questions and identify key issues that will guide a data governance implementation plan, one that balances addressing present versus future needs.

Refocusing a mindset from a singular focus on the IT system to the process of data governance is an important orientation for successful governance.

Data governance sustainability requires that the organization sees data as a strategic asset and that data governance supports the organization's strategic outcomes. For this to happen, key performance indicators (KPIs) for the governance outcomes should be identified, as well as key risk indicators (KRIs) for both the governance outcomes and the operational implementation and management of the data governance.

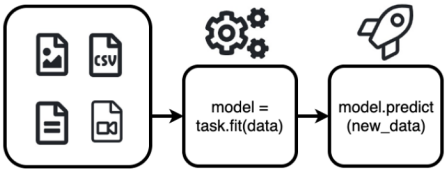
Above all, two overriding factors will influence data governance sustainability. First, credibility and timeliness of the data. Otherwise, the utility of data governance is called into question, and generates frustration and disillusionment in the organization. Second, the development of a data-driven culture must absolutely have all your people onboard, aware and/or skilled (data savvy), and engaged.

About the Author

Betty Ann M. Turpin, Ph.D, President of Turpin Consultants Inc., is a management consultant who has also worked in the federal government, in healthcare institutions, and as a university lecturer. Her career focus is performance measurement, data analytics, evaluation, and research. She is a certified evaluator and coach.

4. Ibid

Amazon Debuted AutoGluon to Democratize Machine Learnings

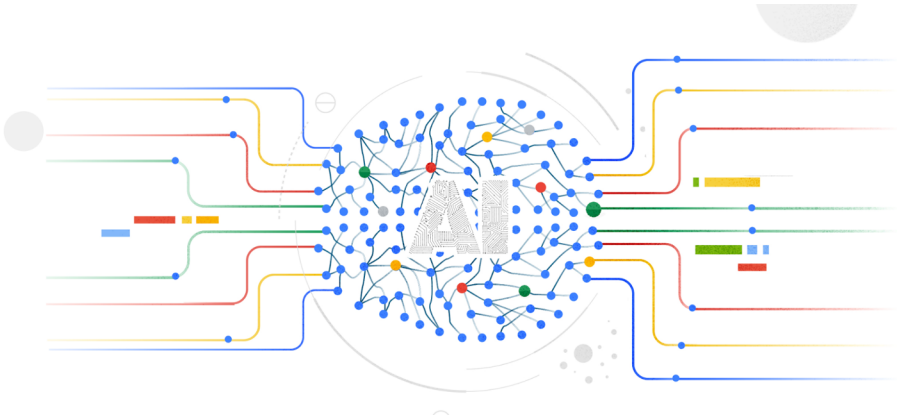


In January, Amazon announced the launch of AutoGluon, a new open-source library for developers building applications involving machine learning with image, text, or tabular data sets. AutoGluon can help developers deploy deep learning models with just a few lines of code.

"We developed AutoGluon to truly democratize machine learning, and make the power of deep learning available to all developers," said Jonas Mueller, AWS applied scientist.

According to Jonas, AutoGluon automates many of decisions that developers have historically had to make themselves, such as hyperparameter tuning and neural architecture search. With AutoGluon, developers can simply specify when they would like to have their trained model ready. In response, AutoGluon leverages the available compute resources to find the strongest model within its allotted run-time. *Source: Amazon Science*

Google Launched Explainable AI Services



Google has added Google Cloud AI Explanations to its cloud platform, which helps humans understand how a machine learning model reaches its conclusions. According to Tracy Frey, Director of Strategy, Cloud AI, the AI Explanations can quantify each data factor's contribution to the output of a machine learning model and help organizations understand why the model made the decisions it did. This tool can be used to improve machine learning models.

However, the tool also has limitations. Tracy explained that AI Explanations reflect the patterns the model found in the data, but they don't reveal any fundamental relationships in the data sample, population, or application. *Source: Google Cloud Blog*

Microsoft launched $40 million "AI for Health" initiative

In January, Microsoft Corp. announced AI for Health, a new $40 million, five-year program, that will leverage artificial intelligence (AI) technology to empower researchers and organizations to address some of the world's toughest challenges in health.

According to the company, the new initiative will focus on implementing AI to accelerate medical research, global health studies and improving access to care for underserved populations.

Source: Microsoft

ANALYTICS IN GOVERNMENT QUARTERLY

ARE YOU INTERESTED IN PUBLISHING AN ARTICLE?

We provide a platform for data enthusiasts to share insights with the public.

Analytics in Government Quarterly

+1 (613) 562-5800 x 5356 | agq@governmentanalytics.institute

ANALYTICS IN GOVERNMENT QUARTERLY

ISSN 2562-9123

FEBRUARY 2020

ISSUE # 2