ANALYTICS IN GOVERNMENT QUARTERLY

FOR GOVERNMENT DECISION MAKERS

FEATURES

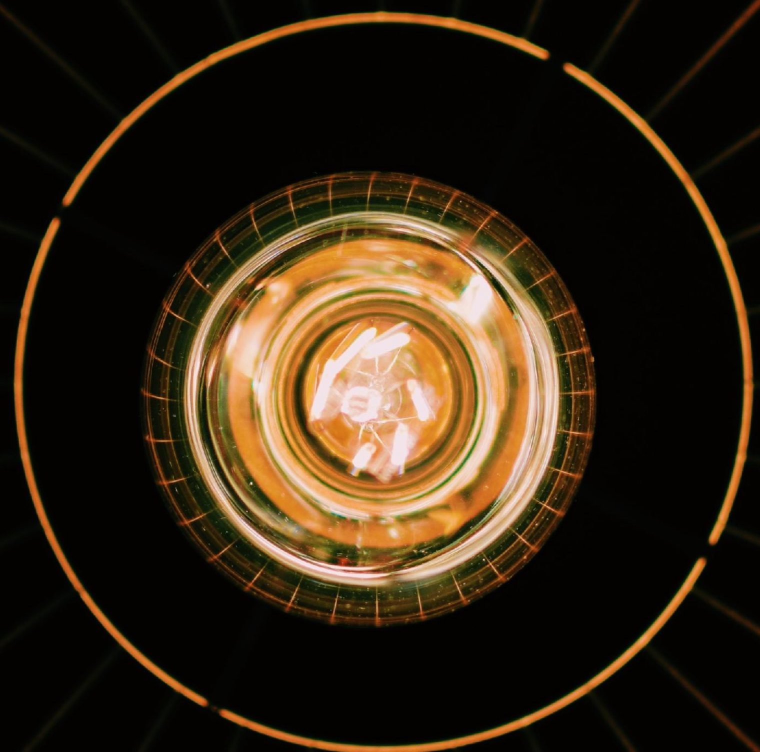
Five Ways towards Responsible Analytics

P. 7

DATA SCIENCE

Getting Data Science Started in Your Organization

P. 16



RESPONSIBLE AI

What is the best way to build fairness, privacy and security while applying AI?

P. 10

ISSN 2562-9123

NOVEMBER 2019

ISSUE # 1

CONTENTS

FEATURES

4/Algorithmic Decision-making: Promises and Challenges for Governments

By Stéphane Gagnon, Alex Ramirez, Gregory Richards

7/ Five Keys towards Responsible Analytics

By Gregory Richards

10/ AI, More Than Just Technology

By Hubert Laferrière and Wassim El-Kass

14/ AI: Risk, Reward and Reality

By Steve Holder & Tara Holland

COLUMNS

17/Point of View

Responsible AI - Are We Getting Ahead?

By Boris Bogatirev

19/Data Science

Getting Data Science Started in Your Organization

By Kevin Kells

22/Dashboard

AI for Decision-Making in Public Administration

By Alex Ramirez

ANALYTICS IN GOVERNMENT QUARTERLY

Editor's Letter

Welcome to this first issue of Analytics in Government Quarterly (AGQ). We are pleased to introduce this journal that will provide insights into research and practical applications of analytics in government organizations. The journal will also become the information distribution tool for the Government Analytics Research Institute (GARI), as well as summarizing research, practical applications, observations, and options from academics and practitioners around the world.

While many other journals and newsletters address the growing field of analytics, AGQ focuses exclusively on government and is meant to provide unbiased, rigorously reviewed and timely information to managers. It will provide practical and useful advice for government managers seeking to deploy analytics of all forms to improve service delivery. On a personal note, having worked in the Canadian federal public service for several years, I have witnessed first-hand the dedication that public servants bring to their work. I have also seen the power that analytics can provide in improving efficiency and effectiveness allowing these highly dedicated people to enhance the value they deliver. It is my hope that this journal provides useful information that can help make this happen.

In this inaugural issue, we address the topic of "responsible" AI and analytics in general. The notion of being responsible for the data we are collecting and the ways in which we use it will become more important as AI initiatives proliferate within government. While we tend to provide our data without much reservation to the likes of Facebook or LinkedIn, there does appear to be some skepticism when it comes to government using our data for purposes that we might not exactly be aware of. In my article on Responsible Analytics, I argue that this problem is not specific to AI; rather, it must be addressed no matter what form of analytics is being used in government organizations. The team from the Immigration, Refugees and Citizenship Canada describes their experience with a pilot project of machine learning for processing visa applications. They point out a basic truth about the use of analytics in government: the technology and algorithms are complicated but well-known and manageable. Another paper written with my colleagues from Carleton University and UQO, argues that the application of analytic tools and automated decision-making calls for close examination of broader governance issues given the potential for litigation related to these types of decisions. Deloitte's AI team provides their perspective on ethics arguing that the use of AI is already prevalent in many organizations, and it is therefore important to explore ethical frameworks that guide practices related to these tools.

We hope you enjoy reading the journal and highly welcome your valuable feedback. Please do contact us here: agq@govenmentanalytics.institute

Dr. 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.

General Inquires:

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

Subscription information:

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

ISSN 2562-9123

Features

Algorithmic Decision-Making:

Promises and Challenges for Governments

By Stéphane Gagnon, Alex Ramirez and Gregory Richards

Government Analytics has become a top priority for public sector organizations at all levels, whether federal, provincial, or municipal. Implementing Artificial Intelligence (AI) technologies and Big Data Analytics (BDA) promises to radically improve performance. As described in other articles in this edition, challenges do exist in integrating the various technologies into the fabric of public sector operations. We present here a brief overview of some promising opportunities and implementation challenges faced by governments in integrating advanced analytics for public programs and services renewal. We address three key issues:

1. Integrating artificial intelligence and big data in operations.
2. Integrating business rules and process automation in programs.
3. Leveraging analytics and intelligent solutions in governance.

Integrating Artificial Intelligence and Big Data in Operations

Government services can be optimized by using AI technologies, especially Machine Learning (ML), but also using Text and Semantic Reasoning, Intelligent Agents, and many others. Implementation is also easier than ever thanks to in-memory Big Data platforms such as Apache Spark and Data Science related ecosystem.

Yet while the "science and technology" is readily available and often open source, many challenges remain to overcome with respect to implementation. Government executives must develop forward-looking analytics strategies, emphasizing realistic yet fast-paced adoption of best practices in bringing ML and other AI technologies to optimize government services.

Integrating Business Rules And Process Automation in Programs

Government programs must ensure constant alignment among 4 key factors:

- The quality expectations of people and organizations served;
- The service standards and processes in place;
- The resource constraints in service delivery; and
- The text of regulations and legislations for programs and standards.

While Business Process Management (BPM) has become well-established, more "intelligent" process automation is now being tested to push even further the "quality-productivity" compromise in the capabilities of government agencies.

It has become particularly essential to develop more open and robust business rules, ensure their integration within processes, and their open sharing among value chain and service ecosystem participants. IT divisions of government agencies must ensure integration of best practices in bringing business rules and process automation in optimizing government programs.

Analytics And Intelligent Solutions in Governance

Algorithmic Decision-Making brings a degree of automation and machine autonomy rarely seen before in government. It implies that we embed in our ML models

and our Business Rules, a whole range of policy insight and knowledge that is expected to function flawlessly in dealing with all cases, allowing agents to handle exceptions with better performance.

However, Public Administration as a discipline has always been focused on conformity, accountability, and assurance of the rule of law and quality standards. The ability of AI and algorithmic government to perform as well as people and teams on these criteria remains to be proven, especially with conflicting prerogatives as to who may ensure the quality of AI and its decisions.

Given recent litigation related to the use of AI in program delivery (e.g., Idaho State in K.W. v. Armstrong) , it is essential to raise key issues that will concern legislators and executives as more advanced technologies are implemented in government programs. Governance bodies, including legislative and judiciary, will need to refocus their attention on complex implications of analytics in public sector, and develop new best practices in modernizing our governance processes and practices, while ensuring that AI remains legitimate and at the service of people and constituencies.

The wide-ranging impact and complexities of Government Analytics make it a most unique and fascinating area for both research and practice. Serving as an agile springboard for innovative best practices and proof-of-concepts, the Government Analytics Research Institute (GARI) invites university researchers, innovation labs in public agencies, and both open-source and proprietary IT vendors to address rapidly-evolving analytics innovation in all areas of public policy and services.

"Algorithmic Decision-Making brings a degree of automation and machine autonomy rarely seen before in government."

About The Authors

Stéphane Gagnon, Ph.D. is an Associate Professor in Business Technology Management (BTM) at the Université du Québec en Outaouais (UQO), based in Gatineau, National Capital Region. He is also a founding member of the Government Analytics Research Institute, a consortium between Carleton University, University of Ottawa, UQO, Institute on Governance and SAS Canada.

Alex Ramirez, **Ph.D**. is an Associate Professor in Information Systems at the Sprott School of Business – Carleton University. He has worked in education for over 30 years. He obtained his Ph.D. from the Molson School of Business – Concordia University in Montreal, a master's degree from Syracuse University in the U.S. and a BSc. High Honours from ITESM, Mexico's top private university.

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.

Five Keys towards Responsible Analytics

How do organizations create responsible analytics strategies?

By Dr. Gregory Richards

"A key point is that many organizations are creating data strategies, but few develop analytic strategies."

The emphasis on responsible artificial intelligence (AI) recognizes that these algorithms can help with decision-making as well as with streamlining certain types of business processes. But because the algorithmic outputs are driven by current data, any bias that exists in the data will also exist in the output. You must have heard of hiring engines that tend to recommend males more than females because the data used accurately reflected the fact that there were more males currently working in certain positions than females. Managers, therefore, should be especially vigilant when seeking to automate or semi-automate decisions.

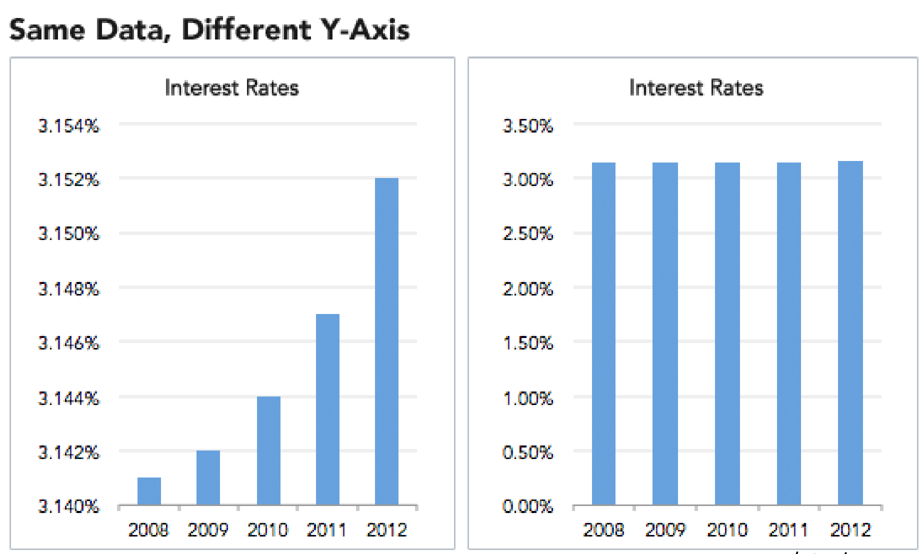
But the problem of "responsibility" is broader than AI—it applies to all analytics being used in organizations. In this paper, I address the broader question of responsible analytics arguing that AI is one analytic tool among many that can be deployed in organizations and that since all forms of analytics use data, managers need to build practices that include responsible analytics.

A key point related to this issue is that many organizations are creating data strategies, but few develop analytic strategies. While the data strategy focuses on sound data management including privacy and confidentiality, an analytics strategy would focus on how data are used in the organization to improve effectiveness. Ensuring responsible use should be part of this strategy.

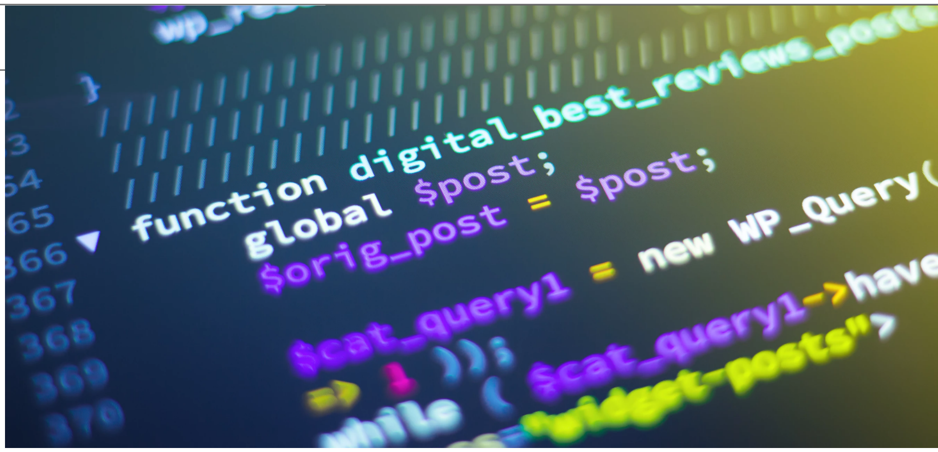
Generally speaking, four components should be integrated into an analytics strategy: visualizations, association modelling, efficiency modelling and what-if scenario planning. Each of these components calls for the use of data in mathematical formulae that lend themselves to manipulation if we don't exercise proper care. This point was made eloquently in Darrell Huff's 1993 book "How to Lie with Statistics". The issues Huff point out are as relevant today as they were in 1993. But the implications are broader given the growing emphasis on analytics in modern day organizations.

Here is an example of a visualization from datapine.com showing how easy it is to mislead with visualizations unless the reader is careful to check the scale on the y axis. The charts fall into the visualization component discussed above. When the graphs are compared side-to-side, it is easy to see that the graph on the left, without a 0 origin, can easily mislead if someone is reading quickly. It's a simple example, but the point is that data in general are used to make decisions. Unless organizations practice responsible analytics, these decisions will lead to some people being disenfranchised depending on the context and the types of decisions being made.

How should managers avoid these problems? One approach is to publish guidelines for visualizations that include proper labelling, consistent axes and required annotations. In other words, being responsible about the use of analytics means creating frameworks for each component in



source: datapine.com



your analytics strategy.

As another example, consider that AI is a sophisticated form in a family of association modeling techniques. One such technique commonly used in organizations is regression modeling, which helps us understand how one thing is related to another. For example, if we want to deliver better services to citizens, a regression model might generate a regression coefficient that tells us that for each new staff person added, service speed will increase by 15 minutes. We use this relationship to plan our investment in staffing.

Regression models depend on the sampling of data from a larger population, and several critical assumptions about the distribution of the data need to be met if we are to have confidence in the output of the model. If these assumptions are violated, we might end up hiring too many or too few staff to meet the demand for services. We exceed our budget on the one hand or drive dissatisfaction on the other. Guidelines, therefore, for checking regression model assumptions should be in place before managers begin to apply these techniques.

In the broader domain of machine learning and AI, an organization's data strategy will often address data quality: accuracy, timeliness and availability of data for example. But, as mentioned earlier, it is possible to use data that accurately represent reality but that lead to biased outcomes.

We should be aware that AI algorithms are based on common mathematical optimization techniques that look for relationships among two or more variables. The algorithm then uses this relationship to predict an outcome. For example, if I'm interested in predicting when someone is likely to retire (what I'll call the "outcome" for sake of discussion), I might gather information on age, salary, number of children etc., on a large number of people. Some of these we know will have retired and some will still be working. I'd then split the data in two to create a "training" set that identifies the regression coefficients and a "test" set that I use to test these coefficients for accuracy. The result is an association model with coefficients (or weights) for some combination of age, salary and number of children that predict who is likely to retire on a new data set where I don't already know the outcome.

The important point is that using a different data set to train the algorithm (i.e., to identify the regression coefficients that predict outcomes we care about) will generate a different association model. Therefore, guidelines about how data are selected, treated and tested along with standards for data accuracy and model validation should be a key feature of analytics use in organizations.

Responsible analytics then, is not just about the protection of data, it's about protection of the rigour and validity of the decisions managers will make using any form of analytic tool. For responsible analytics to prevail, whether in an AI algorithm or some other model, managers need to create the following five conditions:

1. Data sets are representative of the groups for which they will be used.
2. Data sets are as complete as possible and outliers are managed.
3. Data Scientists understand the business problem enough to be able to identify which data sets might not be appropriate.
4. Everyone, Data Scientists and managers, knows enough about responsible analytics to carefully check assumptions before using any form of analytics for important decisions.
5. Guidelines for all components in an organization's analytics strategy are published to ensure appropriate use and deployment of analytic tools and techniques.

AI: More Than Just Technology

By Hubert Laferrière and Wassim El-Kass

A little over two years ago, executive senior management at the department of Immigration, Refugees and Citizenship Canada (IRCC) approved a pilot project proposal submitted by a number of data science practitioners: launching an Advanced Analytics (AA) model to support the processing of visa applications within a business line.

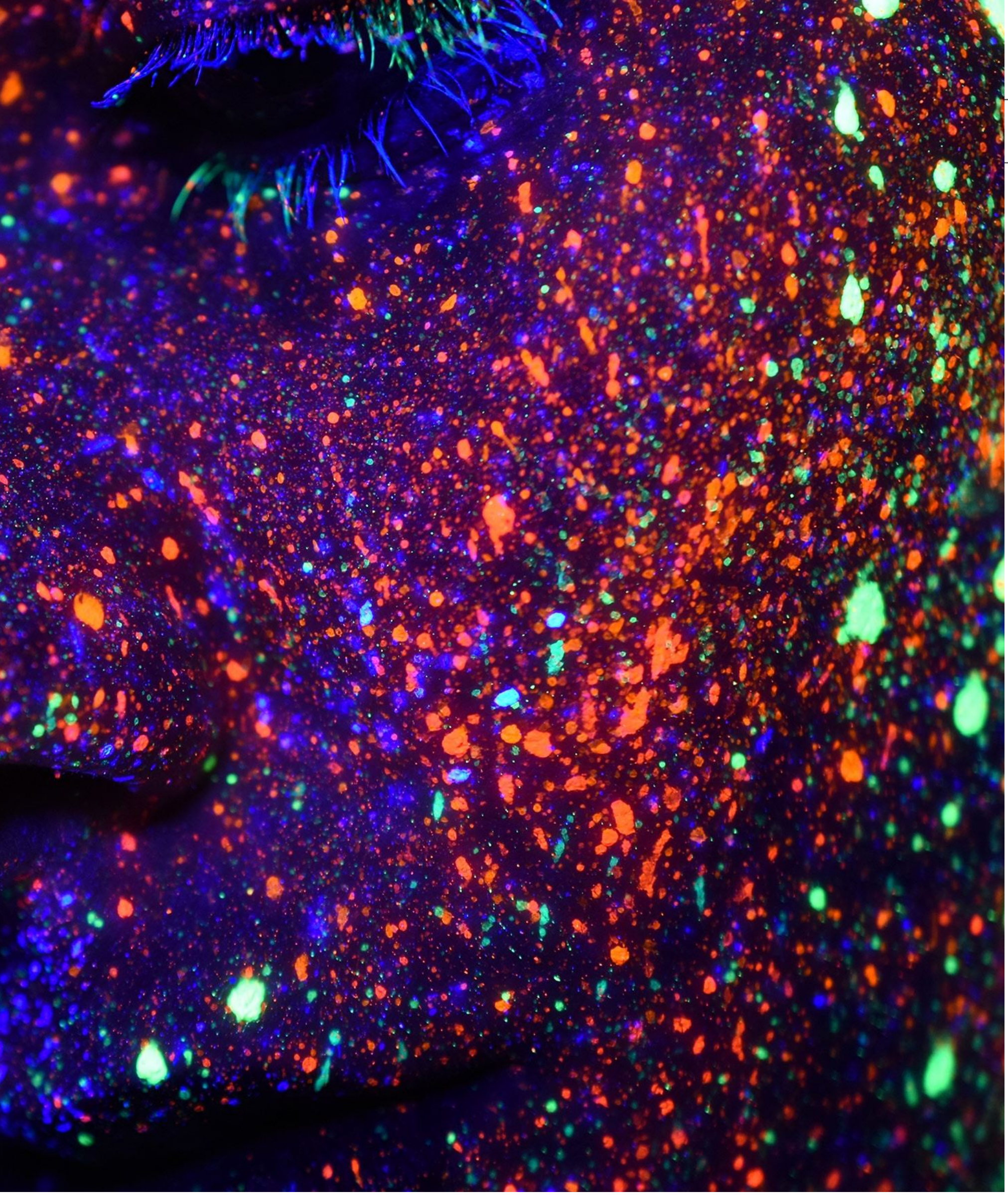
The goal was to help manage the significant and constant volume increase of visa applications using Machine Learning (ML) and Advanced Analytics (AA) technologies. The idea was to triage the applications and automate some activities in the decision-making business process.

The targeted business line was the visa for visitors, in particular those coming from China and India. These countries were chosen since both countries are the main sources of the demand growth and they totaled 50% of the annual volume for visa applications. If successful, processing times could be reduced while tangible benefits could be generated both for applicants and the department.

The positive decision outcome in support of the project was not a surprise to the team. In the weeks preceding the decision, the team had provided some members of the executive management team with results of the analysis and tests we had conducted. They were impressed: it was estimated that up to 35% of the volume of visa applications could have their eligibility determination automatically approved with a confidence level above 99%. Our analysis also showed that by lowering the confidence level by one percent, the volume could increase to nearly 45%. This literally made our Deputy Minister jump out of her chair.

The Challenges Below the water

Once the decision to proceed with the pilot was announced, the team's immediate and main concern was to identify and collect the relevant data, build predictive



models, and deploy them in a production environment as a live pilot. Naively, the team considered that the substantial and significant constraints and obstacles would be mostly technical. While this was partially true, it hid, like an iceberg, the much larger submerged portion.

During the first project management meeting, a plethora of stakeholders knocked on our door raising many interesting questions, potential concerns, and issues to consider. The project team soon realized that the journey was far from being a picnic.

Even though the team knew that the disruptive & challenging technology would raise issues such as the impact on jobs, the team was surprised by their high velocity. The challenges raised were so complex and various that the team questioned whether or not the pilot could be delivered on time or was even feasible.

The challenges were overwhelming due to their diversity, touching every single aspect of the business including program integrity, legal and fairness, privacy, transparency, ethics, labor management, data governance, security, explainable AI, the methodology, and the non-traditional project management. Let's be clear here: most of the key players involved were raising their concerns to look for solutions that would enable the pilot rather than trying to add roadblocks.

AI generates numerous debates in civil society and the "Terminator" image had a profound and symbolic effect in the minds of many. The general public has misconceptions of what AI is and makes incorrect associations with many computer technology failures. Additionally, recent mishaps such as the Phoenix government pay system and the Facebook mishandling of its users' personal information fuel genuine concerns the public may have with IT systems in general and AI in particular when dealing with personal data.

At IRCC, the same concerns and themes had a stronger resonance: as a public organization, fundamental responsibilities and duties were to preserve and maintain procedural fairness of administrative processes and decisions, to ensure transparency and accountability, and to conform to the prescriptions of data protection and privacy laws, rules and directives. It was crucial that all applicants must be equally treated with respect. The Values and Ethics Code for the Public Sector Services are not simply statements on papers, they are a reality by which civil servants must abide.

Dealing with all the issues at the same time stunned the team: for these data science practitioners, algorithms, Advanced Analytics, Machine Learning, and AI in general, were not new technologies per se. The team was well aware that

potential mishaps with harmful effects on citizens could occur. A simple mistake could generate, in a very short time, a significant volume of adverse consequences that would be difficult to repair. Although the odds are different, the risk of this happening exists with any system whether it is using AI or not.

"It is important to strike the right balance between the two while ensuring systems do not introduce unintended bias into decision-making."

Balancing between Human and Machine

The team aimed at resolving the key challenges; one being to achieve the right balance to avoid stifling the development and use of the technology. The team felt sometimes that excessive measures would impede efforts to improve service to the public, optimizing business processes and generating potential savings.

At the outset, our legal team ensured that the right legal framework and authorities were in place for the use and governance of electronic systems, including automated systems. Specific ethical considerations were integrated into the establishment of the ML predictive models: only positive eligibility determinations of applications would be automated; an ongoing quality assurance of these automated determinations would be implemented; the choice of the algorithm had to allow a meaningful explanation of decisions made on client applications (as a consequence, "black box" algorithms, like neural networks, were excluded). The set of rules produced and used by the models had to be approved by the right people.

Experts from the National Research Council of Canada undertook a peer-review: they examined our methodology, approach and models to ensure measures were in place to avoid unintended bias being introduced into decision-making. A Data Readiness Assessment on data quality and governance rules was undertaken. A specialized AI training course on cybersecurity was designed and implemented for data scientists. A Privacy Impact Assessment was completed and measures for enhancing privacy practices with AI were implemented.

Our policy analyst colleagues initiated the publication of a Policy Playbook to support automated decisions. The Playbook outlined guiding principles that would give a coherent basis for strategic choices about whether and how to make responsible use of new tools and techniques. Two fundamental goals were set: (1) the use of new tools should deliver a clear public benefit and (2) humans, not computer systems, are responsible for decisions. The need to curb technology arises, at least as a preventive approach or precautionary principle. The bottom line for us: guide our efforts to ensure human dignity and preserve human values.

Our first challenge is to think about how these systems will be used given their impact on individuals. The use of AA and ML at IRCC relies on a responsible design and implementation approach. Such an approach recognizes the limitations and risks of data-driven technologies and recommends that humans and algorithmic systems play complementary roles. To get the best of each, it is important to strike the right balance between the two while ensuring systems do not introduce unintended bias into decision-making. The analysis of our results shows, so far, that we are striking the right Human-Machine balance.

Generating Substantial Savings

The live pilots were successful. Since April 2018, nearly 140,000 temporary visitor visa applications from China were triaged by our pilot of which 36% were recommended for instant eligibility approval. Ongoing monitoring and quality assurance show that 99.9% of applications recommended for approval by the pilot would have been approved had the eligibility been assessed by an officer. Since the demographics are quite different, only 3% of nearly 143,000 applications received, so far, from India were recommended for instant approval.

We have assessed the impact of models on the business process: when comparing the pre- and post-pilot implementation, we had observed a significant reduction in the time required by officers to process these applications, especially for the instant eligibility approvals. A complete reversal in the proportion of staff working on low-value added versus high-value added tasks occurred: under AA, almost two-thirds of staff were focusing solely on decision-making, rather than completing administrative tasks - tasks that are essential for quality decision-making.

Although it is too early to generalize, initial analysis shows that deploying the models into a full production mode (in 2020) could generate substantial savings to the Department. The Department could, more easily, absorb pressures from the ongoing significant growth of the number of visa applications.

When the official approval of our proposal was announced, the excitement of our data scientist team was most evident. This state of mind remains today but data scientists now know their work entails more than just preparing and using data to train algorithms and build predictive models.

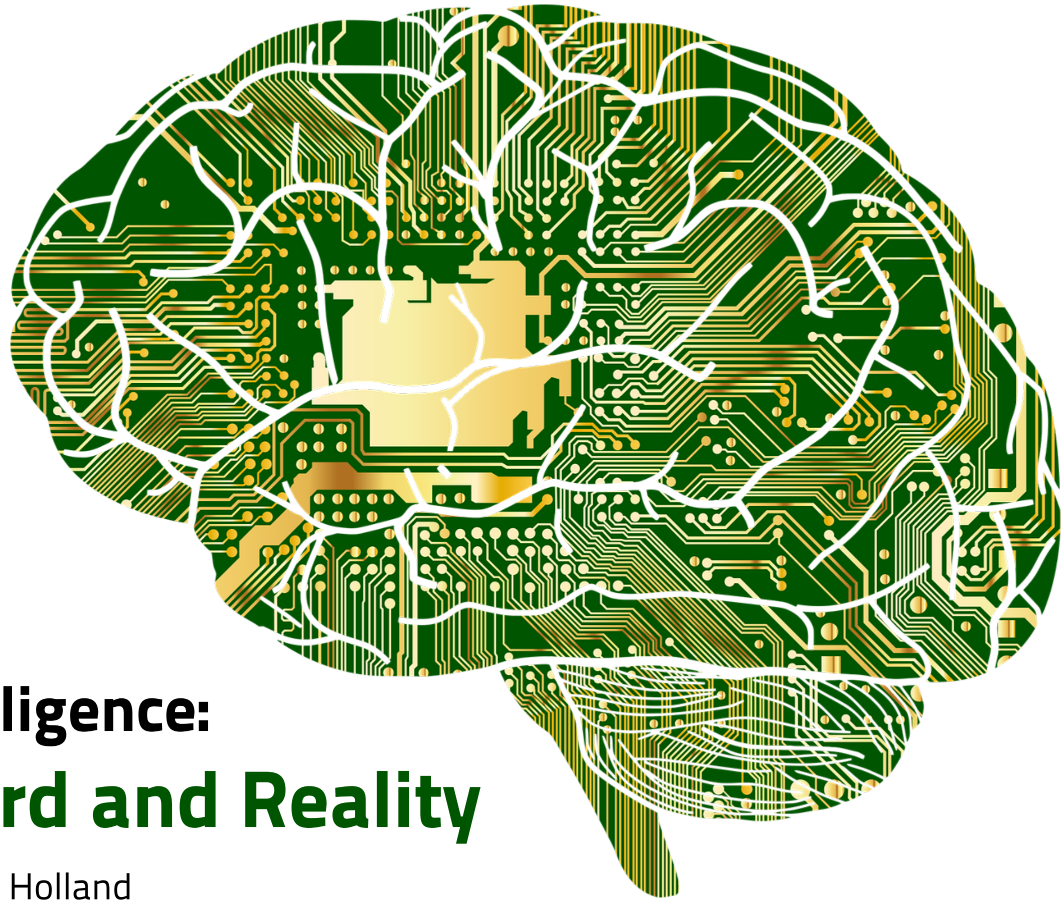
**Special thanks to Steven Gonzalez who managed the data science team during most of the development and deployment phases.*

About The Authors

Hubert Laferrière has established the Advanced Analytics Laboratory for the Department of IRCC. The Lab 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.

Wassim El-Kass is the Assistant Director at the Advanced Analytics Solution Center (A2SC) at IRCC. His 25 years of professional experience ranges from software development to leading development teams and managing projects in both the private and the public sectors. Wassim has a PhD in Information Science and Technology from the University of Quebec.

Features



Artificial Intelligence:
Risk, Reward and Reality

By Steve Holder & Tara Holland

There are a number of perceptions of the impact that artificial intelligence will have on society. Most people focus on the science fiction perception of a sentient robot helping our daily lives or self-driving cars. These are the sexiest applications of AI, the ones that capture the public's imagination. But the majority of AI applications are more pragmatic and address more mundane tasks like finding fraud, assessing risk and predicting or prescribing cause-and-effect relationships in the business.

Ironically, there's much more at stake in the latter. The danger of machines malfunctioning and running rampant á la movies like Terminator or Maximum Overdrive is so remote that it is effectively zero. Meanwhile, real-world consequences—denied credit, immigration standing or lack of access to health care—are real and genuine risks today and have a more direct impact.

It's the so-called "black box" problem. How can we trust and validate decisions made by a machine if we don't understand the algorithms and modeling that make them?

"Canada is taking a leading role in developing governance policies for responsible AI deployment."

The Canadian government is taking the lead in setting governance standards in the application of AI, prescribing a risk-based framework that can be a model for creating an AI-powered organization. The Directive on Automated Decision-Making classifies AI decisions based on the potential impact of their outcomes as well as on the sustainability of ecosystems. The directive makes it clear that AI is not a one-size fits all problem. If an automated decision is going to directly affect the rights, health and economic interests of individuals, communities and entities, the AI application needs to be managed by rules that match the potential harm it could cause. In many cases these rules call for the intervention and review of the decision by humans to ensure appropriate oversight.

These governance standards ensure the Canadian government is doing the right thing for citizens. Level I decisions have minor, easily reversible and brief impacts; Level IV decisions, the most serious, have major, irreversible and perpetual effects. Each level has correspondingly rigorous requirements for notification, explanation, peer review and human intervention. Level I decisions can be made without human intervention and explained by an FAQ page. Level IV systems must be approved by the head of the Treasury Board and require extensive peer review and human intervention at every step of the decision-making process. For more specific information on these classifications and requirements, you can consult Appendix B and Appendix C, respectively, of the directive. The directive is the first of its kind at a national level, and it reflects a commitment by the Canadian public service to ensure a data-driven policy with appropriate human intervention.

Need for Transparency

Transparency is a cornerstone of any customer-facing AI implementation. Users of a service are entitled to understand the process that has an impact on them, whether it's denial of a service, selection for re-assessment, or potentially disruptive land use decision. Processes must not only be fair, they must be seen to be fair.

Not all decisions are equal in impact. The directive's escalating notification scale provides more visibility into the decision-making process according to its impact. And greater visibility into the algorithms and modeling on which decisions are made reveals another paradox of artificial intelligence: algorithmic decisions are more transparent than human decisions. Intuitive decisions are inherently influenced by acquired biases, procedural experience, and fickleness borne of convenience or complacency. At a recent conference on AI in healthcare, one researcher noted that the human brain is, in fact, the black box.

Algorithms can be secure, transparent, free of bias and designed to respect human rights, democratic values and diversity. But they depend on humans providing data sets that are comprehensive, accurate and clean. Users need tools that allow them access across multiple data sets without complex and time-consuming search routines, while maintaining the integrity of that data.

At the same time, that data must be subject to the rigorous privacy standards for which the Canadian government enjoys a well-deserved reputation for leadership. Those data access tools and procedures must have the principles of Canada's Personal Information Protection and Electronic Documents Act (PIPEDA) embedded within them.

Fuelling Innovation

The framework established by the Directive does more than shield Canadians from the arbitrary impact of automated decision-making. It provides a platform for innovation, refinement and bold policy initiative.

At the lower end of the scale, Category I and Category II decisions can be made with little or no human intervention. This is not to say they are not important decisions; the framework assures that the mechanism in place is appropriate to the task at hand. But once those parameters are honed, the more mundane decisions that make up much of our current workload don't demand the attention of a person who can and should be doing more complex work. This frees up program managers and data scientists to ask more keenly crafted questions, prioritize evidence-based policy decisions, and explore possible courses of action in a predictive

and prescriptive fashion. In a sense, the human then is able to take action on the decision rather than crafting the decision in the first place.

This requirement to offload mundane tasks is mirrored in the tools required for such data-intensive innovation. By many estimates, data scientists spend as much as three-quarters of their time cleaning, scrubbing and preparing data for use. Tools that ease the effort and time consumed making data ready to use, free up the human, who has the capacity for curiosity, ingenuity and adaptation for more valuable tasks.

Leveraging AI in Government

Adoption of AI by public service organizations in Canada is uneven, according to research by SAS, Accenture and Intel. Pockets of government have robust and well-governed AI capabilities, whereas other organizations haven't entered the AI discussion.

As of April 1, 2020, new policy requirements come into play that require peer review not just for AI outputs, but for AI projects themselves. But there are no guidelines for conducting those reviews. IOG, in partnership with GARI, has been approached by several departments to be the facilitator and convenor of peer reviews, and develop guidance for the Treasury Board Secretariat regarding the peer review mechanism.

To make these policies more comprehensive, the TBS must explicitly define these processes and provide ready-made tools to support the Directive, both internally and in citizen-facing engagements. There must be dialogue with industry experts to resolve the "black box" issue, and recognize the AI inherently supports the goal of efficient, accurate, consistent and interpretable decision-making and transparent governance.

Private Sector Applications

While the directive is designed for government decision-making, it can also serve as guidance for private sector. Private sector firms can also benefit from appropriate guidelines for applications of AI, machine learning, natural language processing, neural networks and other automated decision-making technologies. As in the public sector, business applications of AI have a range of consequences: an inappropriate purchase suggestion from an online storefront does not have the same impact on a user's life as the denial of a mortgage. The tiered system of the directive provides ample room for application regardless of the use case.

It's a useful thought experiment to envision a near-future AI-enabled application—say, for example, real-time insurance rate adjustment—and categorize in according to the directive's tiers. What is adequate documentation, a referral to an FAQ page or real-time e-mail notification? What is the duration of the impact? How difficult would it be to remediate the impact of a faulty outcome? The tiered system accommodates interpretation and a variety of appetites for risk, and is open to evolution as new applications are conceived and real-world environments change.

As the public sector embraces evidence-based policy and businesses deliver new models for serving their customers and shareholders, AI-augmented decision-making will serve an ever-growing role. The Directive on Automated Decision-Making provides an opportunity for Canadian organizations, public and private, to take bold steps forward in this emerging frontier.

About the Authors

Steve Holder, the Head of Strategy and Innovation at SAS Canada, is responsible for growing revenue and building the long-term vision for SAS in the Canada market. He owns the SAS solution strategy, modernizing the SAS' customer base and managing the innovation ecosystem including higher education alignment. Steve is the SAS Canada evangelist for the technology including Artificial Intelligence, Cloud and emerging technologies.

Tara Holland currently leads the Canadian Federal Government Specialist team for SAS Canada, spearheading Public Sector efforts by providing data & analytic expertise and business guidance to government and healthcare organizations across the country. With over 20 years supporting government customers her passion is in bridging the gap between Business and IT and enabling organizations to deliver on the value of AI.

By Boris Bogatirev

Responsible AI – Are We Getting Ahead?

The acceleration of the fourth industrial revolution paired with accessible and almost endless computing resources (as Gordon Moore had predicted) is shining a strong light on Artificial Intelligence (AI). With a plethora of movies, shows and books to feed our wild imagination on "how and when the machines will take over", it is natural to talk about ethics and governance. Personal privacy is an ever-growing concern that continues to generate news and controversies on a daily basis. Margins are shrinking and commercial competition is growing stronger. All of these factors contribute to us relying more and more on automation to process vast amounts of data to produce these coveted insights that would help our organizations win. With any progress comes the natural question of what is the price that we are paying in exchange.

There are multiple questions that we should be asking regarding responsible applications of AI:

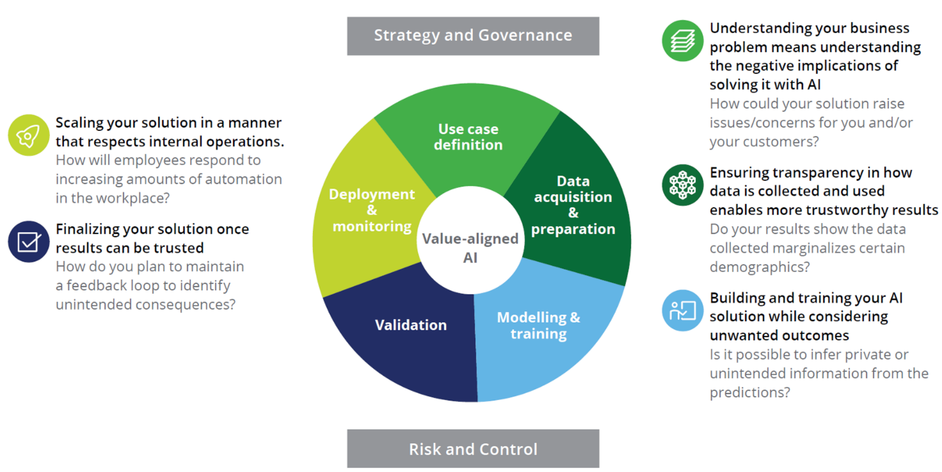
- How can we (humans) make sure that the machines are not planning to take over?
- Will my role in the workforce become obsolete because of AI?
- Will AI algorithms make fair and unbiased decisions about consumers and citizens like me?
- How can I trust an AI output if I do not understand how it works?
- Will my personal information be used by AI algorithms without my consent?

AI is expected to be one of the leading economic drivers of our time, and Canada has the opportunity and responsibility to be a global leader. As a country, we have the research strength, talent pool, and startups to capitalize, but that is not enough if we want to lead in an AI-driven world and shape what it might look like. True leadership which means taking steps now to establish a world-class AI ecosystem in Canada, is required.

AI is no longer on the horizon. It is here now, and is already having a profound impact on how we live, work, and do business. In fact, most statistical methods behind various AI algorithms have been around for decades. The questions citizens and consumers may have concerning AI means that businesses need to consider the ethical implications and underlying risks throughout the life cycle of an AI application, and to have a clear strategy on how to evaluate and balance the risks and benefits of implementing AI. Executives will face challenging decisions about how AI applications should be built, what values should be upheld, and whether they should be built in the first place. While ethics are contextual and the perception of them depends in large part on geography, culture, social norms, organizational values and more, they cannot be ignored. Companies need to be intentional in the way they address the numerous ethical implications of AI solutions and how they will respond to unintended consequences.

Ethics, both implicit and explicit, have a role at every stage of the AI life cycle. Everyone involved in AI development must be responsible for identifying and responding to ethical concerns. As organizations accelerate their adoption of AI technologies, they must address the various ethical questions that can arise throughout the entire development life cycle. As an example, most organizations are working to ensure data privacy is protected and datasets are unbiased, but these challenges are just the tip of the iceberg. Ethical issues are often missed because of ambiguity about what ethics mean and the lack of accountability about addressing them. For example, consider the following ethical dilemma for an autonomous car: if it must hit a person, should the autonomous car hit a child or an elderly person? In these cases, the way forward is unclear and not everyone agrees as many moral frameworks could apply. Questions related to AI's purpose and values, e.g, where and how it should be used, fall into this category, as do questions about the future nature of collaboration between AI and humans in the workplace. These are longer-term challenges that will likely require new forms of collaboration and discussion to address them. At the same time, this is not the first time that technology has hit up against profound questions.

The following framework can be used to consider ethical and responsible applications to help organizations adopt AI and to use it responsibly.



The management of AI ethics cannot be a periodic, point-in-time exercise. It requires continuous support and ongoing monitoring. Without a holistic perspective of the AI life cycle and the different moral and organizational barriers AI could present, organizations can open themselves to fundamental threats to their operations. Identifying and categorizing ethical issues is one of the main challenges an organization will face concerning AI. Understanding the range of concerns is the first step to addressing and mitigating the problem.

The bottom line? Organizations need to decide how they will identify and manage the ethical considerations around AI if they want to benefit from these technologies and ensure their long-term success.

**Four industrial revolutions are: water and steam to mechanize production, electric power to create mass production, electronics and information technology to automate production respectively, and digital revolution respectively.*

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 Fortune 500 companies.

By Kevin Kells

Getting Data Science Started in Your Organization

"I know the status of data science being used in my department. And I know where we're headed. We have a realistic plan, taking gradual steps towards fully utilizing the data we produce. We are using data in ways to inform our entire organization from management down to guidance at street level for our caseworkers. Our employees are informed about it, feel engaged, and are excited about where we're going. They're happy to take part, especially those who have shown interest and curiosity and who have opted in to the online learning modules that we've made available to them. And those who are less technically inclined know they will benefit, too, without having to learn programming or statistics. Our stakeholders in the bigger picture are also satisfied with the multi-layered approach we're taking to respect and protect the privacy and the confidentiality of the data we manage."

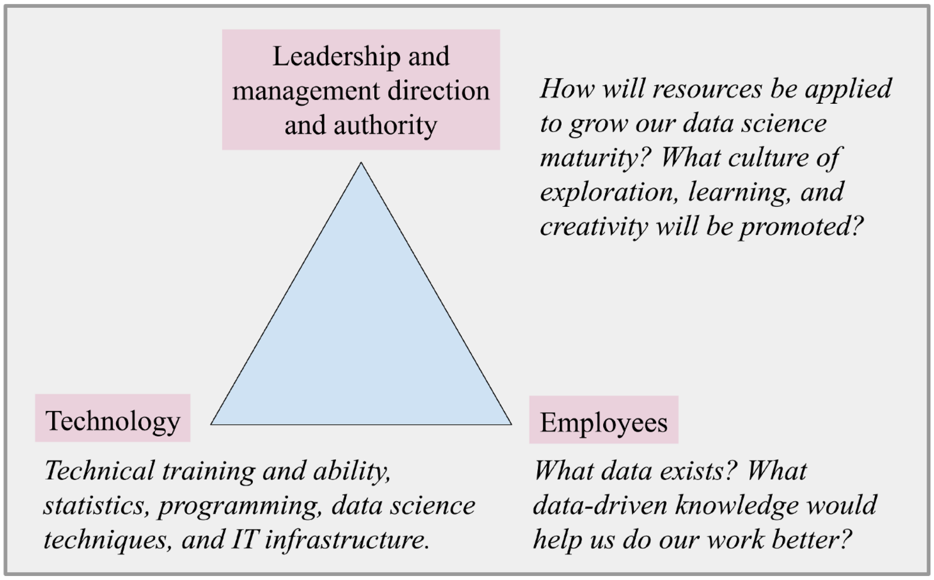
Does this sound too good to be true? If so, it may be a relief to know that you are not alone. There are many organizations who also think the aspirations expressed in this statement are beyond the reality they are experiencing. They are hindered not just by available resources — both budgetary and human — but by a lack of clarity in big-picture understanding of what data science can be doing to help their organization and what they can be doing to grow their organization's "data science maturity". If your organization's most advanced data science is — or until recently was — limited to Microsoft Excel, you are in good company.

The adoption of data science in an organization evades the linear principles of other changes, for example a change in technology tools rolled out by the IT department, a change in how business is done as driven by management, or a change in culture, as exemplified and promoted by leadership. The reason is that the adoption of data science involves a change in all three of these aspects at the same time. The introduction of new technology tools is just one of three aspects in data science adoption. The integration of data science into business processes and thinking is another aspect that involves those doing the work identifying how data can be used, and then adopting the use of that data in workflows and operations. Finally, there is a requirement in data science for a measure of openness to trial and error, a sense of exploration, and a touch of creativity.

These three aspects of data science adoption are depicted in the following graphic below. Here, the three aspects are laid out clearly. Let's start by looking at the least important aspect, technology. This aspect involves the planning and installation of IT systems that can be used to store data, and software that can be used to analyze data and visually and interactively present results as information and knowledge. This

"Data science maturity is a matter of a culture of learning, employee participation and technology. In that order of importance."

technology aspect also includes the creation (by training and education) or importation (by hiring or outsourcing to consulting firms) of the knowledge in statistics, programming, and data science techniques required to properly transform data into meaningful and useful information and knowledge.



The next aspect involves the employees. It is not through technology, nor through leadership and management, that we answer these questions: 1) What data do we have, and how can we leverage it? 2) Given the available data, what kind of knowledge and information can be extracted that would relate to the daily needs of the work experience, be valuable to improving the quality and efficiency of work results and the quality of work life?

The role of employee engagement, participation and input is key to answering these questions, and we'll give some suggestions shortly about how to go about cultivating this input, no matter in what stage of data science maturity your organization finds itself.

Finally, the role of a culture of learning and exploration is a key concept to grasp, especially for the leadership and management, who may mistakenly consider the growth of data science maturity in the organization to be a matter of simply purchasing IT systems or outsourcing number crunching to a consulting firm. Why is learning the most important? Because it is through learning, exploration, and creativity that your employees — your organization — discovers what data is available that it hadn't previously considered, how to extract knowledge and information from that data, and how that knowledge and information can be integrated into the daily experience, the operations, and the workflows throughout the organization. The factor of fun in enticing employees to engage in identification of data and how to use it should not be underrated. In fact, the concept of "gamification" in data science is gaining more and more recognition. It has little to do with "games" and much more to do with unlocking creativity, exploration, and discovery, and engaging end-users—your employees—in a way that for them is enjoyable and intellectually stimulating, and yet effective. If the participation of the employee is key to discovering an organization's data and its most effective use, it is a culture of learning and exploration promoted by leadership, which will empower employees to participate.

What can an organization do?

The curious chasm can be bridged between management's desire to grow an organization's data science maturity and the role of the individual employee's exploration and personal experience in growing that maturity — complicated by the third factor of technology — software, hardware, and knowledge of statistics, programming, and data science techniques. Here are three action items to consider.

1. Promote an ecosystem for data science exploration among employees. Even without knowing which employees will step forward with ideas, let your organization know that leadership is deciding that there is likely latent data science curiosity, if not talent, and that managers, in an atmosphere of curiosity and discovery, should keep their eyes out to cultivate interest.

2. Invite employees at any level to take introductory online courses in data science such as edX online courses (www.edx.org) organized

by University of California Berkeley, M.I.T., Harvard, and other highly regarded universities. Your organization can sponsor a number of seats and reimburse employees for the cost, which, e.g., for UC Berkeley's "Data 8X: Foundations of Data Science" on edX costs around US$ 270 per student for a 2-4 month course. Keep track of which of your employees sign up and complete the course, as they will be the first ones to follow up with to invite their ideas for how to apply what they've learned in the daily work of their teams.

3. Sponsor periodic data science summits within your organization, where employees or employee teams can showcase to the entire organization efforts and results from data science projects they've worked on within their departments. Winning recognition and even prizes for the best data science projects provides the rare benefit of moving all three aspects of the needle at the same time. First, it encourages employees to explore what data is available to them and what information or knowledge would be most meaningful for them to extract to incorporate in their work experience. Second, it highlights, to those wishing to get started, who in the organization has technical knowledge and technical resources and how to contact them to share and export that technology and those technology ideas to other parts of the organization. Finally, leadership sponsorship of data science summits within the organization sends an unmistakable message that exploration is encouraged and even rewarded.

"The biggest challenge of making the evolution from a knowing culture to a learning culture—from a culture that largely depends on heuristics in decision-making to a culture that is much more objective and data driven and embraces the power of data and technology—is really not the cost. Initially, it largely ends up being imagination and inertia."

—**Murli Buluswar**, *Chief Science Officer, AIG*[1]

An organization's data science maturity relies on the balance of a three-part effort to progress. Invite your employees to explore and to propose which data they think is useful for them and how they'd like to see it analyzed and visualized. Support the employees with education and with the technology they request to work on the projects they propose that incorporate data science into their workflows. Reinforce the culture of discovery and exploration; data science talent among existing employees can be nurtured and their efforts to explore projects can be enabled and shared within the organization. This will not just inspire and empower other employees to learn and explore, it will create an atmosphere and ecosystem for data science discovery that will be attractive for hiring data-science-ready new talent.

Reference:

[1. www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/how-companies-are-using-big-data-and-analytics](https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/how-companies-are-using-big-data-and-analytics)

About the Author

Prof. Kevin Kells 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 as well as M.Sc. and Bachelor's from Georgia Tech in Electrical Engineering.

Kevin 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. Kevin's research interests lie in approaches to large scale, complex human challenges at the confluence of Government, Industry, Academia, and Non-Profit sectors; and in how to improve the use of data analytics in government and non-profit organizations for the greater good.

By Alex Ramirez

AI for Decision-Making in Public Administration

In order to balance power and accountability, public administrators engage in decision-making, analyze their policies and the inputs that have developed such policies and those needed to produce alternative ones. It is precisely in these two aspects, decision-making and analysis, that artificial intelligence can be an incredible resource. Working in the field of decision support systems for the last 20 years has allowed me to see the field evolve, from Decision Support Systems (DSS), Group Decision Support Systems (GDSS), Business Intelligence (BI), to Business Analytics (BA).

In all these solutions, the main driver of the changes has been technology, but it has been fuelled by a constant environment of change. Organizations are under constant pressure to make better decisions. They have adopted technology to respond to or anticipate three main events: Problems, Opportunities, and Directives.

Problems

Among the many factors affecting organizations, some are considered problems. These need to be addressed as soon as possible. A problem is recognized when the organization is unable to achieve their goals and objectives. In order to recognize the problem, the organization must have ways to measure its performance and compare it against its goals and objectives. Many do this through scorecards and different types of reports. This must be a daily activity to avoid surprises when it is too late. Some factor-generated problems are globalization, customer demands, market conditions, competition, etc.

Opportunities

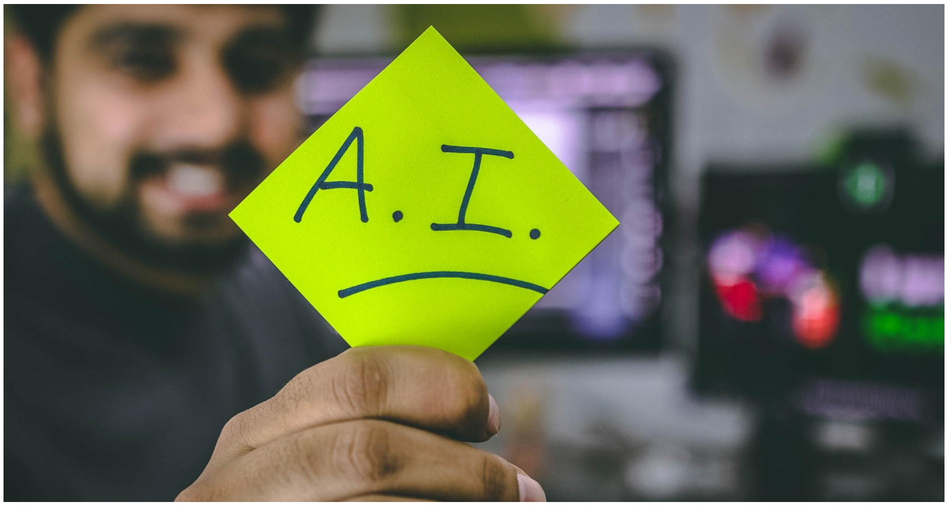
Organizations that keep up to date on their environment, can observe or anticipate emerging trends that have not been explored, and they see these as opportunities. Not all these opportunities are worth pursuing, but in order to assess them, they need to have some data to measure their potential impact. Opportunities that are considered viable, must be pursued as soon as possible, before someone else sees them. In other words, the race to be the first to cash in on them is the fuel behind these events.

Factors that give emergence to opportunities include new technologies, changes in market conditions, news about competitors' struggles, trade wars, etc.

Directives

Every organization is accountable to its stakeholders and is regulated. Managers are free to act as long as they meet their board's expectations, follow their regulations and do not break the laws. But sometimes, they need to accept changes in their operations when new regulations are put in effect. These directives force them to react in a way that will generate the least amount of disruption in their daily activities. Think for example what happened when the Harper government dropped the GST in July 2006 and then again in January 2008. All Canadian organizations of every size had to comply with these regulations and modify their systems accordingly.

What all these events, problems, opportunities and directives, have in common is that to understand them, organizations need data to measure and address them. Organizations can respond by:



changing strategy, looking for ways to increase productivity, generating new business models, looking for collaborations, creating new products, or modifying their supply chains. Each of these responses generate some options. These options must be evaluated. The best way to evaluate them is using a cost-benefit analysis (CBA). For each option, its costs must be quantified. This is an easy task since costs are always an upfront issue. A more difficult exercise is to quantify its benefits. Benefits are always a long-term issue and forecasting is not an easy task. Once data for both costs and benefits are collected, managers can engage in CBA, and determine the impact of each option. Only then, will they be able to make decisions. Decision-making is only possible when there are viable alternatives and each alternative is properly assessed.

The amount of resources needed to engage in decision-making increases with any additional option considered. This is perhaps the reason most organizations have adopted computer systems to aid them in the decision-making process. If all the data needed for evaluating these options is available digitally, there are several tools that will shorten the time needed for their analyses, giving managers the opportunity to engage in What-If analyses, where they can anticipate the impact of each one of these options and have a better grasp on them.

What-if analysis helps managers understand how the solutions are impacted when an input variable is changed. If data are available and a tool is able to calculate the impact of the changes using a set of models that will answer immediately any of the questions posed by managers, what if instead of 10% reduction we negotiate a 12.5%? is it viable to produce 500 additional items? Do we have enough raw material? How fast can it be acquired?

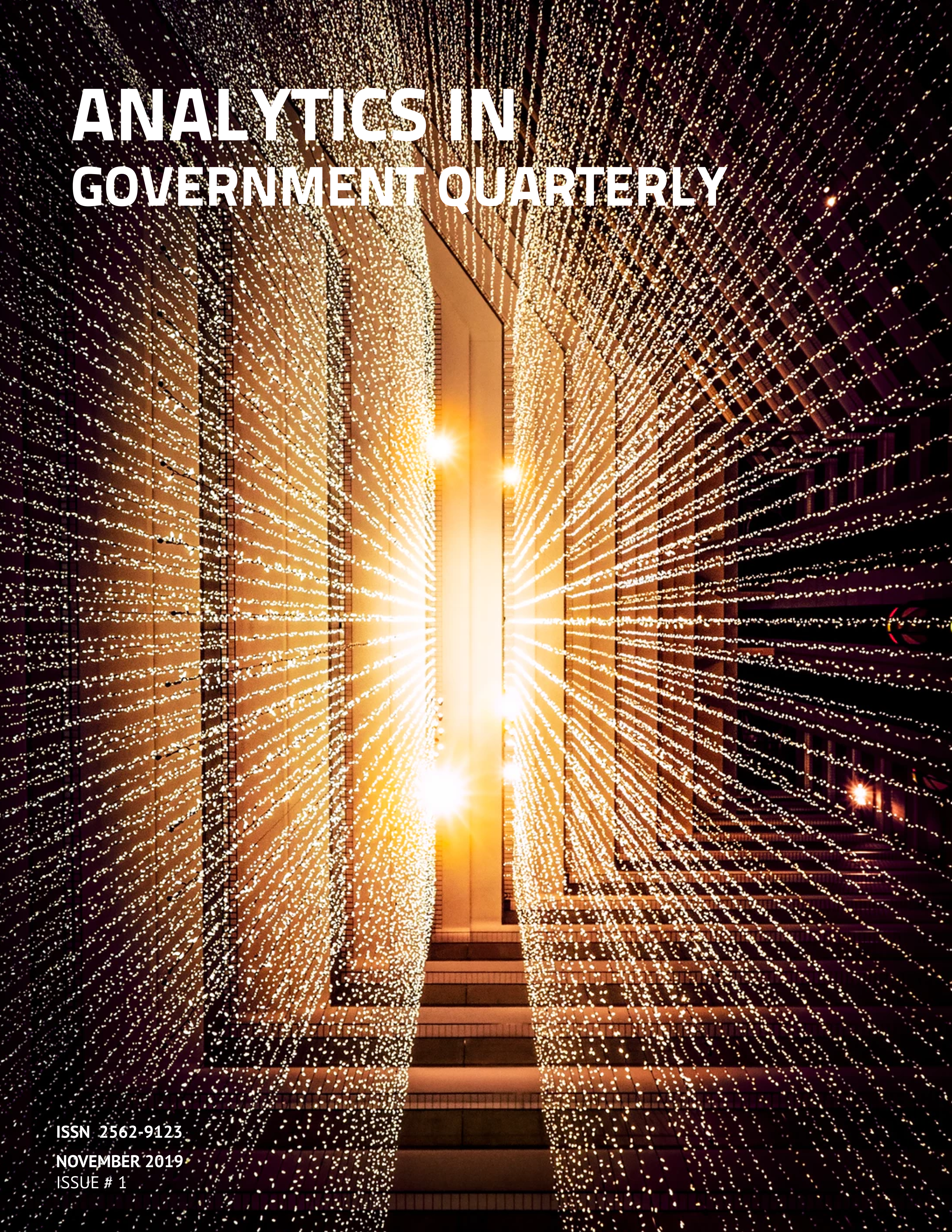
Artificial Intelligence (AI) solutions can now be incorporated in the process of decision-making to help managers ask questions that previously have not been considered. These AI solutions can be added to the models used and help managers in the event of facing a problem, seizing an opportunity and/or following a directive.

Therefore, if public administrators who want to respond to the constant environment of relentless change and still balance power and accountability, they ought to improve their decision-making processes. They can do so by learning how to use AI models that will give them additional insight when analyzing their policies. These models also can help them identify what additional data can help in capturing the inputs used when developing new policies, and how the data can be justified to propose alternative ones.

At all levels of government, we need public administrators that are knowledgeable and proficient in the use of systems that can help them in their decision-making activities and that are able to use AI algorithms to their advantage to justify and become more accountable for the impact of their policies. The issue is no longer whether to use them, but how and how soon, either to solve problems, seize opportunities, or implement directives.

About the Author

Alex Ramirez, **Ph.D**. is an Associate Professor in Information Systems at the Sprott School of Business – Carleton University. He has worked in education for over 30 years. He obtained his Ph.D. from the Molson School of Business – Concordia University in Montreal, a master's degree from Syracuse University in the U.S. and a BSc. High Honours from ITESM, Mexico's top private university.



ANALYTICS IN GOVERNMENT QUARTERLY

ISSN 2562-9123

NOVEMBER 2019

ISSUE # 1