

**Canada Revenue Agency
Business Intelligence Strategy
October 1, 2014 to March 31, 2017**

A strategy for data-centric innovation



Canada Revenue
Agency

Agence du revenu
du Canada

Canada

TABLE OF CONTENTS

BI Strategy at a Glance 1

Introduction 2

BI at the CRA – The Journey So Far..... 5

BI Landscape for Tax Administration..... 7

Corporate Strategies and Initiatives 8

Applying BI in CRA Programs and Services..... 14

Conclusion 16

BI Strategy at a Glance

STRATEGIES	PRIORITY INITIATIVES
CORPORATE AGENDA	
THEME: STRONG FOUNDATION	
The knowledge, skills and capacity to deliver, support and leverage the BI environment and the right mix of processes, end user tools and enabling technologies to efficiently supply and use reliable and secure integrated data.	
Recruit and develop people with the right knowledge and skills	<ul style="list-style-type: none"> Talent management program Partnerships with academics
Establish the capacity to act quickly	<ul style="list-style-type: none"> BI tools and infrastructure renewal Data/analytics sandbox Data program led by Chief Data Officer
THEME: IMPROVE INTEGRATION	
Increase capacity to conduct research and analytics projects that span programs and the organization.	
Increase opportunities and efficiency through a whole-of-Agency perspective	<ul style="list-style-type: none"> Advanced analytics unit Behavioral economics (Nudge) unit
Increase availability and timeliness of access to quality, integrated data	<ul style="list-style-type: none"> Redesign the Managed Metadata Environment Data integration Phase II BI data usage framework for internal and external data
PROGRAMS AGENDA	
THEME: SUPPORT PROGRAM AND SERVICE DELIVERY	
Plan, prioritize and execute program-focused BI projects based on a rigorous assessment of impact on Agency costs and performance.	
Apply analytics and behavioural economics techniques to support the CRA's compliance, service and integrity objectives	<p><i>Initiatives vary by program and sponsoring branch. Some examples are:</i></p> <ul style="list-style-type: none"> Explore holistic (cross-program) risk score Advance risk-based approaches in audit and debt management Use Nudge campaigns to increase reporting and payment compliance and uptake of e-services

Introduction

Business intelligence (BI) is information derived from the data available to an organization. It is about gleaning knowledge and insight from that data. BI is not a new area of interest for the Canada Revenue Agency. We collect and create vast amounts of data as we conduct the activities required to administer taxes and benefits for Canadians, and to internally manage our organization. We have been using this data to derive information about program performance for many years.

Big Data and BI Infrastructure

Big data is frequently described in terms of three characteristics:

- volume, the amount of data in terms of consumed storage;
- variety, which refers to a mix of structured, unstructured and semi-structured formats; and
- velocity, which refers to the speed at which the data are created or captured.

The McKinsey Global Institute, in its report *Big data: The next frontier for innovation, competition, and productivity*, offers a more pragmatic “beyond the ability of typical database software tools to capture, store, manage, and analyze”. This link between the big data concept and the technologies that support its processing is important, for it is our ability to use the data that is the key differentiator. For example, people have been clicking on web pages since the last century, but the potential insights to be gleaned from the web traffic remained hidden until the data could be captured and used to understand the business implications.

The notion of “beyond typical database software tools” is also of note. Enterprises do not have to adopt entirely new sets of technologies to take advantage of BI, but they must be willing to recognize when the limitations of the existing tools and infrastructure are reached and specialized technologies are required.

Globally, BI is in a period of resurgence. We are seeing massive increases in the amount of digital data created every day, from the web, social media, sensors, and other sources. Technology is evolving to enable the capture and processing of these data. (See *Big Data and BI Infrastructure sidebar*.) Most importantly, there has been a realization in the business environment that there is a potentially significant payoff to be had if these data could be tapped to improve products, services and/or customer reach. These factors are combining to push BI well beyond its reporting and performance measurement roots.

This is particularly true in the private sector where many recent success stories have originated. By making fundamental changes to their business processes through the use of advanced data analytics, these organizations are seeing that the payoffs for BI investments are making measurable impacts on the bottom line. For example:

- Three major U.S. banks – Wells Fargo, Bank of America, and Discover – are using analytics to understand aspects of the customer relationship, particularly multi-channel interactions that occur through websites, call centres, direct dealings with bank representatives and regular banking transactions;

- In retail, companies are capturing and analyzing the consumer’s “digital footprint” and using the results to drive customer loyalty through the provision of customized special offers and other enticements; and
- Credit card companies are combating fraud by identifying, in real time, transactions that don’t fit the cardholder’s usual spending patterns.

Analytics and Behavioural Economics

While the term analytics can refer to any logical analysis, in the BI context it refers to the more advanced forms of data analysis, such as statistical analysis, forecasting, and predictive modeling. While much could be written about each of these techniques, they share a common characteristic: they go beyond simply reporting what happened to providing insights about why something happened, what if the trend continues, and what will happen next.

The potential value to be derived from their application, for example to support timely evidence-based decisions, is significant.

Behavioural Economics (BE) is a very specialized field that integrates a psychological perspective into analyzing environment and motivating factors in order to predict and influence (“nudge”) behaviours toward a desired outcome. More specifically, a Nudge is a policy intervention that targets the environment in which individuals make decisions that have important implications for their well-being. Subtle features of decision-making context are re-arranged or added to direct behavior toward the best option. In many cases, a Nudge intervention can be a cost-effective way to facilitate voluntary compliance with desired outcomes.

Public sector organizations are recognizing that similar techniques can be applied to their own business challenges. Advancements being made by other tax administrations are particularly relevant to CRA. For example, HM Revenue & Customs (HMRC), the UK tax agency, has invested in BI solutions to capture and analyze multiple internal and third party data sources to discern relationships and patterns. These are then applied in their work to tackle fraud and evasion.

The UK has also been a leader in the development and application of behavioural economics techniques. Its Behavioural Insights Team (BIT) has worked with a number of UK ministries and now includes the provision of advice to other governments as part of its mission. In 2013, the White House announced the establishment of a behavioural insights team for the U.S. (See *Analytics and Behavioural Economics sidebar.*)

One of the key factors influencing the pace of adoption of advanced analytics and behavioural economics techniques is the need for specialized skillsets. The acquisition, development and retention of talent are challenges the Strategy needs to address. In 2011, McKinsey stated¹ “a significant constraint on realizing value from big data will be a shortage of talent, particularly of people with deep expertise in statistics and machine learning, and the managers and analysts who know how to operate

companies by using insights from big data”. Three years later², “talent challenges are stimulating innovative approaches—but more is needed. Talent is a hot issue for

¹ **Big data: The next frontier for innovation, competition, and productivity**, McKinsey Global Institute, June 2011

² **Views from the front lines of the data-analytics revolution**, McKinsey Quarterly, March 2014

everyone. It extends far beyond the notoriously short supply of IT and analytics professionals ... The management and retention of these special individuals requires changes in mind-set and culture”.

As the number and variety of available data sources grows, data consumers – and particularly governments – must remain vigilant to protect privacy. The CRA can learn from the knowledge and experiences of private enterprise and public entities in other jurisdictions, but we must apply these lessons learned in a Canadian tax administration context.

Perhaps most importantly, we must be prepared to move forward with the knowledge that we will not necessarily get it 100% right the first time. As one participant in Gartner’s “business analytics in the enterprise” field research³ noted: “True analytic innovation comes from trying new things”.

This business intelligence strategy – the CRA’s first – seeks to put the Agency on a path to data-centric innovation.

³ **Why Business Analytics Projects Succeed: Voices From the Field**, Carlie J. Idoine, Gartner, September 12, 2013

BI at the CRA – The Journey So Far

The CRA collects and creates a significant amount of data in carrying out its mandate to administer tax and benefit programs for the Government of Canada and most provinces and territories. This data – supplemented with limited amounts of data from other government organizations – is available for BI use, subject to the same rigorous oversight and strong controls that are in place to secure and protect any other use of personal or confidential information.

CRA began its BI journey in the realm of reporting, and these first generation BI solutions – which are used to meet CRA and GC reporting obligations, and for program quality assurance and planning – still represent a large portion of the Agency's BI capacity. The skills and technologies to capture and use data from Agency operational systems into the Agency Data Warehouse and numerous Data Marts are mature.

Analytics are also being used in program areas to look at past taxpayer behaviour to better understand future compliance risks. Today, risk-based approaches have transformed the audit, debt management and non-filer workflows. It is now a common practice to assess risk and use risk scores to inform workload selection and intervention approaches, allowing workload to be directed to areas with the highest risk of non-compliance and potential for tax recovery, resulting in measurable savings.

To date the use of data mining models has provided gross annual benefits of approximately \$180M. – TSDMB

By analyzing operational results to understand the impact of the compliance interventions, the CRA is able to continuously improve the management and development of its compliance programs and activities.

Business Intelligence permeates our business processes; everybody uses BI every day, so we need to be leaders in this area. – CPB

The Agency also maintains a specialized workforce that provides professional advice, statistical services and analytical support to clients for statistical, tax, fiscal and socio-economic policy making. Stakeholders and clients include other federal departments and agencies, the provinces and territories, the public and private sectors, international bodies and many areas within the CRA. In 2012, a BI Centre of Expertise was

established in the Strategy and Integration Branch to further develop the BI capability and to provide expert advice on approaches to research, analytics, data mining techniques and methodology.

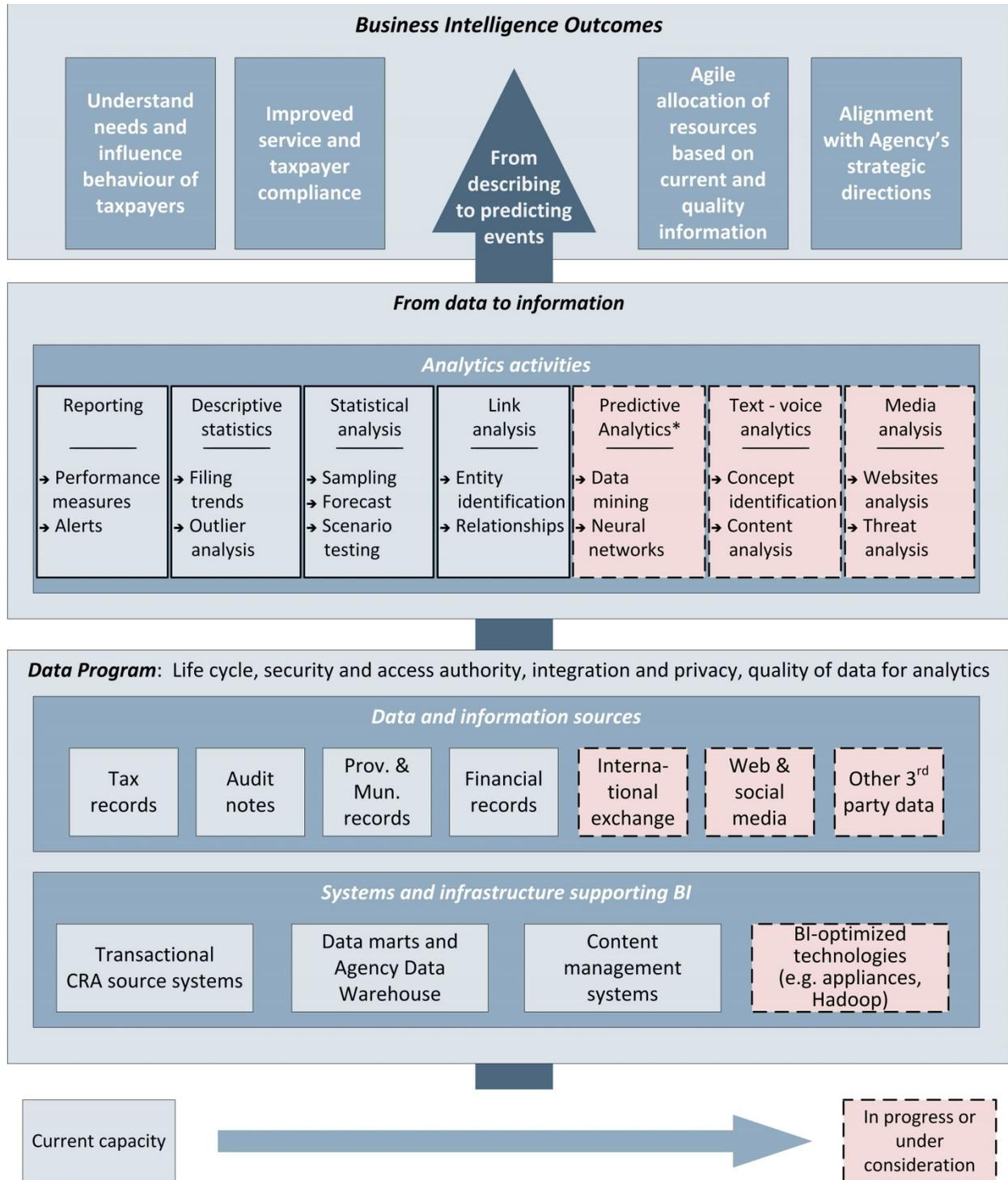
CRA's specialized workforce of statisticians, economists, methodologists, and program officers produce a variety of data products addressing all business lines.

The CRA is ready to take the next steps on its BI journey, to position the Agency to further leverage the power of data to advance our strategic goals. In order to take full advantage of the inherent opportunities, the Agency needs a BI capability at the ready. We will seek out collaboration and partnership opportunities to maximize the value of our investments, and to apply discovered insights across programs, to the full scope of service and compliance interventions.

We expect that the right application of BI techniques will allow us to:

- Deliver results more efficiently, such as maximizing tax recovered through our compliance activities by reducing the number of “no change” audits;
- Be more agile in addressing emerging risks and challenges, such as those presented by the underground economy; and
- Be more responsive to taxpayer needs and expectations, and offer them the services they need through their preferred channels.

BI Landscape for Tax Administration



*Predictive analytics capacity is already in place in the Taxpayer Services and Debt Management Branch

Corporate Strategies and Initiatives

STRONG FOUNDATION		
Initiative	Year 1 (Oct. 2014 – Mar. 2016)	Year 2 (Apr. 2016 – Mar. 2017)
BI talent management program	<ul style="list-style-type: none"> Establish guidelines for the use of short term (assignment-based) external resources Create a recruitment program, including entry-level development opportunities 	<ul style="list-style-type: none"> Create BI Talent Management program management office Develop career path and initiate joint internal processes Launch the recruitment campaign
Partnerships with academics	<ul style="list-style-type: none"> Establish a framework for academic partnerships Engage academics in at least one Nudge and one Big Data project Develop data-partnership guidelines 	<ul style="list-style-type: none"> Initiate at least one tax-relevant partnership research project based on CRA data Create networks to facilitate graduate-level involvement in the Agency
BI tools and infrastructure renewal	<ul style="list-style-type: none"> Conduct planning and analysis 	<ul style="list-style-type: none"> Continue with tools/infrastructure analysis and design. Implementation schedule to be coordinated with Shared Services Canada.
Data Analytics Sandbox	<ul style="list-style-type: none"> Procure and install the necessary technologies, and make ready for use Test the sandbox with one research or analytics project 	<ul style="list-style-type: none"> Measure value and make recommendations for ongoing use
Data program / Chief Data Officer	<ul style="list-style-type: none"> Propose roles and responsibilities of Chief Data Officer Propose responsibilities and placement of data program 	<ul style="list-style-type: none"> Establish Chief Data Officer Establish CRA data program Issue first report from the Chief Data Officer on the state of the Agency's data

People – Addressing the Skills Gap

To take full advantage of data, particularly for research that crosses program and organizational boundaries, the Agency needs to build and sustain a workforce adept in advanced data analysis, possessing a blend of business, mathematics, and computer science expertise. For Nudge applications, specialized knowledge in behavioural economics is also necessary.

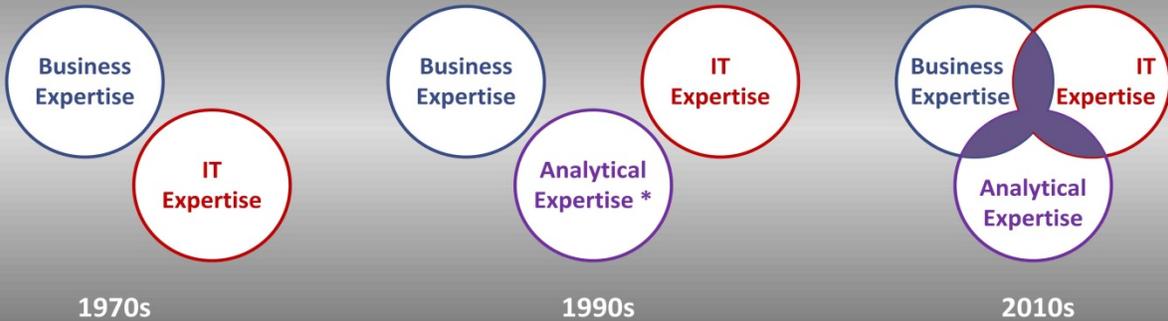
Even with a strong talent pool, it will be difficult for the Agency to keep up with the rapid advancements in BI techniques and their application in different business scenarios. As well, there may be times when the Agency requires specialized skills that it does not possess in-house, or additional capacity that is only needed for a short duration, such as for a particular research project. Formalized, long term relationships with academics

in fields relevant to CRA can bring a range of unique world views, expertise and innovation to the Agency.

TALENT MANAGEMENT

“A significant constraint on realizing value from big data will be a shortage of talent, particularly of people with deep expertise in statistics and machine learning, and the managers and analysts who know how to operate companies by using insights from big data.”
(McKinsey Global Institute)

SKILLS REQUIRED FOR A SUCCESSFUL ENTERPRISE

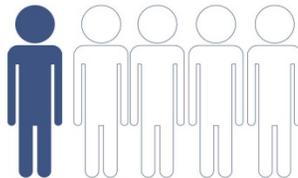


* The skillsets within the scope of *analytical expertise* continue to evolve and now include data mining, predictive and behavioural analytics, econometrics and forecasting, in addition to more traditional mathematics, statistics and research disciplines.



New roles are emerging that bring together business, IT and analytical expertise. But as with anything new, the supply of these resources does not align to the demand.

Data scientists, and others with deep analytical skills, are the discoverers of data insights.



Projected shortage in the United States by 2018: 140,000 – 190,000 individuals.

Data-savvy business managers and analysts are the consumers of data insights.



Projected shortage in the United States, across all business sectors, by 2018: 1.5 MILLION individuals.

Educational programs are emerging, but are not expected to keep up with the demand in the short term.

Established and planned offerings at Canadian post-secondary institutions:

- 🇨🇦 Professional Masters in Big Data, Simon Fraser University
- 🇨🇦 Collaborative Master's Program in Data Science, Carleton University
- 🇨🇦 MBA in Business Analytics, Sprott School of Business
- 🇨🇦 Certificate in Data Analytics, Big Data, and Predictive Analytics, Ryerson University
- 🇨🇦 Certificate in Management of Enterprise Data Analytics (Big Data), University of Toronto

Tools and Technology – Optimizing the Data Environment

Analysts and researchers must be able to execute their BI activities in an environment that performs efficiently and can support the demands placed on it. Specialized technology can improve performance in the supply and use of data. To optimize delivery times, the Agency must also determine which types of BI activities will continue to require support from IT personnel, and which can be effectively set up on a self-serve basis by data analysts and researchers using established tools and capabilities.

While the standard BI platform and processes, and the solutions built upon them, will meet most needs, we expect that there will be some research and analysis that is best carried out in a dedicated “sandbox” environment. This sandbox would be characterized by very fast set-up in bringing diverse cross-program and potentially external data together for use, flexibility in the types of data that can be included (structured, unstructured, other files), and high processing speed. This type of environment is not suitable for all BI users. Specialized knowledge in both data and business would be a prerequisite for use, but this sandbox would be highly effective in supporting the kind of one-off, high-profile key questions to which the Agency regularly is called upon to respond.

Strategic Data Leadership

Currently at the CRA, no single business entity has responsibility for the Agency’s overall data holdings. Differences in accountabilities affect the Agency’s ability to derive maximum value from data in its possession or that may be available externally. Additionally, a whole-of-Agency perspective – rather than an approach closely tied to the delivery of particular solutions – can improve how the agency collects, uses, manages, publishes and shares data.

The CRA is proposing that a data program, under the leadership of a Chief Data Officer, be established to provide a business-led coordinated approach to the acquisition, governance and use of data.⁴

⁴ The data program is not exclusively a BI need. The Chief Data Officer would also be expected to drive the Agency’s open data agenda, aligning with Government of Canada Open Data directions. However, the program and leadership role are expected to provide benefits in the BI space as well.

DATA LEADERSHIP & GOVERNANCE

“... the executive suite needs someone who can oversee the strategic business application of its information assets enterprise-wide.”

Micheline Casey, Chief Data Officer (CDO), Federal Reserve Board



In 2013, 100 large organizations employed CDOs, more than double the number in 2012.

Gartner predicts that by 2015, 25 percent of large global organizations will have appointed CDOs

Top 3 industries adopting the CDO role:

**Banking
Government
Insurance**

Distribution of CDOs:

**65% - United States
20% - United Kingdom
15% - Other (12 countries)**

IMPROVE INTEGRATION		
Initiative	Year 1 (Oct. 2014 – Mar. 2016)	Year 2 (Apr. 2016 – Mar. 2017)
Advanced Analytics Unit	<ul style="list-style-type: none"> Define team make-up and staff team in SIB Create lab environment for team use Work with program branches on selected analytics projects 	<ul style="list-style-type: none"> Identify an Agency transformation opportunity for the team to work on
Behavioural Economics (Nudge) Unit	<ul style="list-style-type: none"> Create Behavioural Economics (Nudge) team Work with program areas to identify and carry out candidate Nudge initiatives Create guidelines for Nudge development and implementation process 	<ul style="list-style-type: none"> Deploy resources to support priority initiatives across Agency
Redesign the Managed Metadata Environment	<ul style="list-style-type: none"> Conduct analysis and design for an enhanced metadata registry/repository 	<ul style="list-style-type: none"> Continue with design and construction. After the strategy period, CRA will transition from the existing MME to the new enhanced tool.
Data Integration Phase II	<ul style="list-style-type: none"> Set scope and parameters for Phase II Conduct analysis Conduct pilot 	<ul style="list-style-type: none"> Make recommendations on long term integration solution(s)
BI Data Usage Framework	<ul style="list-style-type: none"> Draft BI data usage framework 	<ul style="list-style-type: none"> Finalize BI data usage framework

Centralized analytics and behavioural economics capacity

To better respond to projects with whole-of-agency scope and strategic importance, the creation of a small centralized team of analytics experts is proposed. The team will undertake high-profile data and analysis work in support of Agency and government priorities, supported by program-based expertise on a project-by-project basis. A centralized team – operating under a *hub-and-spoke* model that provides services and expertise across the Agency with resources that can be deployed to specific assignments on a short-term basis – will also increase the CRA's agility for program areas that have only occasional demands for these skills, for whom developing and maintaining their own capacity is not cost-effective.

Internationally and within the Government of Canada, there has recently been a strong interest in using insights and approaches gleaned from different disciplines such as Behavioural Economics (BE), psychology, and sociology to inform decision making and policy development related to a wide range of compliance issues. Identifying candidate Nudge projects, and designing the most appropriate intervention, requires specialized expertise. The *hub-and-spoke* model recommended for the advanced analytics team is

equally applicable to the development and deployment of a behavioural economics capability at the CRA. In collaboration with program experts, this centre of expertise will research data model designs, objectives and sampling methodologies to develop a range of effective and targeted Nudge strategies to increase taxpayer compliance and service offerings that are based on key program priorities.

Data Provisioning and Use

The capture and use of business and technical metadata enhances data quality and makes it explicit, an important consideration when data is used for BI. The Managed Metadata Environment (MME) is a CRA-developed registry and repository that currently houses limited business metadata, along with a variety of technical metadata. The redesign of the MME was recommended to enable the capture and use of additional business metadata properties.

In addition to metadata, integrated data is fundamental to successful BI, particularly when tackling complex business problems or research questions. A recurring requirement at the CRA is for a complete view of a taxpayer, over time and across multiple program interactions. The first phase of the Data Integration Feasibility Study was conducted between 2012 and early 2014. The second phase will use a 'big data' approach to integrate a greater volume and variety of data, seeking to minimize labour-intensive manual preparation. The goal is to develop a methodology for integrating data across business lines for both individual and business tax filers, and to understand the processes and technology needed to implement this integration into an enterprise reporting system.

To ensure risks associated with data usage are properly identified and managed, we use a variety of tools – such as Privacy Impact Assessments and Threat Risk Assessments – during the planning and preparatory phases of Agency projects. However, these tools were not designed with BI data usage in mind and additional factors need to be considered. For example, business intelligence data sets may be comprised of very large amounts of data, but the data may be anonymized to eliminate or mask identifying characteristics, or they may include external (non-CRA) data. Assessing the risks associated with using the data, and determining the appropriate controls to mitigate the risks, needs to be carried out in a BI context.

A BI Data Usage Framework, aligned to the CRA Integrity Framework, to govern and facilitate BI data usage will be developed to ensure that the CRA continues to act with the utmost of integrity in the handling of taxpayer and personal data.

Applying BI in CRA Programs and Services

Some specific opportunities have already been identified to apply one or more advanced BI techniques within and across programs during the planning period. These are highlighted below. More plans will be formulated during the strategy period. While the use of advanced analytics and behavioural economics techniques are a focal point for the Agency's BI efforts over the next several years, all programs do not need to adopt these capabilities to the same extent or at the same pace. For some programs, continuing to evolve their reporting capabilities may be the BI technique most suited to their needs now and for the immediate future.

Program area	Year 1 and Year 2 (Oct. 2014 – Mar. 2017)
Enterprise	<ul style="list-style-type: none"> • Explore the development of a holistic risk score for tax entities that can be used across programs
Audit	<ul style="list-style-type: none"> • Expand the scope and refine the application of behavioural economics (Nudge) techniques to reduce the number of cases requiring a full audit • Share BI results across programs/branches to optimize decisions, for example to better understand the linkages between activities in ABSB, TSDMB, Appeals, and CPB • Conduct analysis on the portability and predictability of risk scores across audit programs
Debt Management	<ul style="list-style-type: none"> • Expand the use of the TSDMB integrated research environment to continue to improve the risk-based approach to debt management program delivery • Make use of behavioural economics (Nudge) to influence taxpayer debt payment behaviour and maximize the use of program resources • Recruit, train and develop a debt management workforce with knowledge, skills and expertise in BI • Strengthen partnerships with other organizations to identify and exploit new ways of detecting and addressing non-compliant behaviour related to filing and payment
E-Services	<ul style="list-style-type: none"> • Develop Nudge campaigns to increase the level of e-payments, explore the delivery of electronic Notices of Assessment through the Agency's secure portals, and to continue to increase the e-filing uptake

Expanding the Use of Risk-based Approaches in CRA Programs

Building on the advancements already made through the use of risk scoring to support the Agency's compliance activities, the CRA will explore the development of a holistic risk score that considers a complete view of the taxpayer and could be used across programs.

In the Audit program, CPB has identified a need to expand its risk view beyond audit-specific financial risk to include behavioural elements and risk indicators from other

programs and other parts of the Agency, to more of a whole-of-taxpayer level. Doing so will allow the Agency to move from yes/no decisions on audit to a model where based on risk the Agency will select from a range of interventions – audit, review, information visit, promotion or communication, encouragement of self-reassessment, Voluntary Disclosure Program suggestion, investigation or no action.

In the Debt Management program, TSDMB plans to continue using advanced analytics in combination with behavioural insights and sophisticated segmentation to separate the compliant from the non-compliant taxpayers. This will allow them to apply the appropriate tax debt treatment ranging from soft treatments such as service, education, and communication, through to targeted compliance actions including the use of legal enforcement tools.

Leveraging Behavioural Economics Approaches

Increasingly, workloads within the compliance programs are managed by selecting from a range of interventions, based on the factors relevant to a particular case. As the Agency's behavioural economics expertise grows, Nudge approaches based on these principles will be considered along with other more traditional interventions.

On the service side, ABSB plans to use BI – and in particular behavioural economics – to gain a better understanding of taxpayer needs and behaviour, to use that insight to help design the right services, and to move taxpayers to e-services.

Conclusion

Whether focused on improving tax and benefit program outcomes, or looking inwards to our administrative programs, the possibilities for BI are numerous. The Strategy and Integration Branch will continue to play a leadership and advocacy role, working with all parts of the Agency to promote business intelligence, and advance our collective knowledge, maturity and use of BI techniques.

The initiatives in this Strategy are expected to be funded through a combination of new investments and existing base budgets. Requirements will be specified in the detailed work plans.

Strategy progress will be monitored using two complementary measurement tools:

- **Agency-level BI measurement** – Strategy progress will be reported annually to senior management.
- **Initiative-level BI measurement** – Detailed plans for each initiative will specify the objectives and measures to be used, as well as the governance and reporting model.

Through this BI Strategy – the Agency's first – we will continue to apply BI techniques in our programs – and across programs – to transform approaches and improve outcomes. Through our corporate initiatives to further strengthen the CRA's BI foundation and improve integration, we will position ourselves to take advantage of new and ever-changing opportunities. And as we do so, we will take stock of the results so that we can learn from them, and apply the evidence-based insights we gather to future challenges.

We will become an Agency that uses data to fuel innovation.