Appendix: Tables

Table 1 Interviews conducted

Interview Format	Field	Identifier	Interview Date
Oral Interview	Federal and Provincial Public Service	Public Servant A	2013
		Public Servant B	2014
		Public Servant C	2013
		Public Servant D	2014
		Public Servant E	2014
		Public Servant F	2014
		Public Servant G	2014
		Public Servant H	2014
		Public Servant I	2014
		Public Servant J	2015
		Public Servant K	2015
		Public Servant L	2015
		Public Servant M	2015
		Public Servant N	2015
	Consultant	Consultant A	2013
		Expert	2014
		Consultant B	2014
		Consultant C	2013
		Consultant D	2014
	Non Profit	NP A	2015
		NP B	2015
Written Response	Public Service	Treasury Board Secretariat	2014
		Communications Security Establishment	2014

Table 2 Access to information requests

Dept. or Agency	Status
Privy Council Office	Complete, Sept. 2016
Privy Council Office (No.2)	Complete, Sept. 2016
Canadian Heritage	Complete, Aug. 2016
Statistics Canada	Complete, Sept. 2016
Global Affairs	Complete, Feb. 2017
Treasury Board Secretariat	Complete, Nov. 2016
Industry Canada	Extension, not yet received
Finance Canada	Complete, Oct. 2016
National Defence	Extension, not yet received
Employment and Social Development Canada	Extension, not yet received
Department of Justice	Complete, Aug. 2016
Canada Revenue Agency	Complete, No Documents exist
Shared Services Canada	Complete, July 2016
Public Safety	Complete, Dec. 2016
Citizenship and Immigration	Complete, July 2016
Canada Border Services Agency	Complete, No Documents exist
Health Canada	Extension, not yet received

Table 3, Title of documents received as part of access to information responses

Dept. or Agency	Document Title and Type	Date
Privy Council Office (PCO) (23 pgs.)	DM Standing Committee on Big Data 'Leveraging Big Data for Internal Government Use', presentation to DM Committee by Industry Canada DM Committee on Policy Innovation Agenda	Nov. 2014 Nov. 2014 Nov. 2014
Canadian Heritage (CH) (209 pgs.)	Correspondence Policy Research Group Contribution to Fact Sheet Fact Sheet: DM Committee Meeting on Economic Trends and Policy 2016 Departmental Environmental Scan 2016-2017 Departmental Research Plan Understanding Big Data: Policy Research Group Proposed Contracted Public Opinion Research Project 2016-2017	2014-2016 2016 Nov. 2014 Dec. 2016 Feb 2016 May 2016 2016
Statistics Canada (SC) (421 pgs.)	Project Plan: CPI Alternate Data Sources Project Charter: CPI Alternate Data Sources A Suggested Framework for the Quality of Big Data Big Data at Statistics Canada, 2013-2014: A Progress Report	Feb. 2015 Oct. 2014 Dec. 2014 March 2014

	Diagnostic Report: Steering Committee on Big Data	March 2015
	Big Data for Health Statistics Division	Sept. 2015
	Walkability and the Canadian Community Health Survey	Unknown
	Creating a Framework to Add Environmental Data to the CCHS	March 2016
	Creating a Framework to Add Environmental Data to the CCHS	March 2016
	Big Data Status Report	May 2016
	Big Ideas Conference Boot Camp Proof of Concept Statistics Canada Integrated Health App	Unknown
	Adding Spatial Data to Maximize the Analytical Potential of the Canadian Health Measures Survey	Unknown
	Survey Data, Administrative Database, and Walkability Linkage: An Update	2014
	Big Data Survey 2015 – UN Global Working Group on Big Data	2015
Global Affairs (GA) (208 pgs.)	Occasional paper: Big data and DFATD	2016
	Big Data and Global Affairs Canada	2016
	Global Affairs Canada: Big Data	2016
	Business Intelligence Competency Centre: Project Kick Off (Gartner Consulting Presentation)	2016
	Direct Diplomacy Moving Forward 2015-2016	2016
	Open Source Research Processes and IT Risks	2016
	Direct Diplomacy Activities	2014
	The Future of Direct Diplomacy	Unknown
	Direct Diplomacy: Open Source Data Driving Foreign Policy	2014
	Direct Diplomacy: Supporting Canada's Democracy Agenda	2014
	Results-based Management for Direct Diplomacy Direct Diplomacy	2015
	Direct Diplomacy	Nov. 2015
	Direct Diplomacy: Supporting Canada's Democracy Agenda	2015 Unknown
	Direct Diplomacy: Use of social media in advancing democracy and protecting and promoting human rights	Unknown
	Direct Diplomacy: Understanding non-state actors in the digital space through big data tools	2015
	Request for Information: Current Analytics Practices	2015
	DFATD Web Statistics for 2015	2016
	GAC Web Stats 2016 to the Month of May	Unknown
	Presenting Corporate Data to the Public Anticipatory Media Lines	2016
Treasury Board Secretariat (TBS) (124 pgs.)	Email Correspondence	2014
	IC Presentation to DM Committee on Policy Innovation	Nov. 2014
	Steering Committee on Big Data Diagnostic Report	2015
	Presentation preparation worksheet	2015
	Leveraging Big Data for Internal Government Use, Presentation to the Committee on Information Management in Business	Feb. 2015
	Diagnostic Report, Steering Committee on Big Data	Dec. 2014
	Service and GC2.0 Tools Policy and Community	Dec. 2015

	Enablement	
Finance Canada (FC) (11 pgs.)	Note: Economic Action Plan Summary for New Brunswick	2014
Shared Services Canada (SSC) (85 pgs.)	Big Data @ SSC presentation Big Data – Session 2 De-brief of Big Data Discussion from Architecture Framework De-brief Big Data – Architecture Framework Advisory Committee Big Data – Session 3	June 2015 Sept. 2015 Aug. 2015 Dec. 2015 Jan. 2015
Justice (261 pgs.)	Correspondence Big Data at Justice and Privacy Implications overview Possible Big Data Uses by the Department of Justice (three drafts) Big Data Workshop, Draft Agenda Key Activity Summary Memo to Senior Assistant Deputy Minister, approval for workshop Memo to Deputy Minister, Outcomes of Big Data Research and Strategic Planning Workshop Big Data Strategic Planning Workshop Nov. 25, 2015: Outcomes Report	2015-2016 July (2) and Sept. 2015 Nov. 2015 August 2015 Nov. 2015 2016 Unknown 2015
Public Safety (PS) (26 pgs.)	'Finding the Balance: Big Data Analytics in the World of Safety and Security'. Presentation to the Director General Portfolio Policy Committee	April 2015
Citizenship and Immigration Canada (CIC) (43 pgs.)	Strategic Communications Plan Social Media Campaign Evaluation Framework	Unknown

Table 4, Access to information requests obtained that were filed by others

Department or Agency	Document	Date Document Produced
Privacy Commissioner (PC) (39 pgs.)	Briefing Note: Big Data and Privacy Briefing Note: To summarize Obama administration white paper 2012 Briefing Note: Summary of the interim progress report on 'Big Data: Seizing Opportunities, Preserving Values' Briefing Note: Meeting with Industry Canada Minister James Moore Briefing Note: Industry Canada's Digital Strategy Summary of Standing Committee on Industry, Science and Technology Meeting	May 2014 March 2012 Feb. 2015 Feb. 2015 April 2014 No date
Canada Revenue Agency	Business Intelligence Strategy, Oct. 1 2014 to March 31	2014

(CRA) (18 pgs.)	2017	

Table 6, Mapping big data applications across Government of Canada

Dept. or Agency*	Activity
Government of Canada	Suggested the Government could use data collected through the Internet of Things for fleet management, the military, police, health care, citizen satisfaction, service delivery, and employee production. No details provided (SSC 2016). The Government of Canada is researching how it may develop and use a public service cloud.
National Resources Canada	Exploring social media analytics to tailor communications; developing automated visual processing to support manufacturing clients; analysing data from Canadian Vehicle Use Study to develop fuel consumption policy and program initiatives; involved in data generation and collection through its Federal Geospatial Platform (TBS 2016, p. 84).
Environment Canada (Environment and Climate Change Canada)	Makes use of sensors through its work with the Meteorological Services of Canada
Health Canada	Exploring how it can use big data to improve the way it monitors disease, air quality, and the safety of imported products; genomics research; projecting healthcare expenditures and financing options for health care (TBS 2016, p. 12); First Nations and Inuit health surveillance, health promotion, addictions, claims processing and health benefits (SC 2016, p. 150).
The Genomics Research and Development Initiative	to develop new treatments for chronic and infectious diseases, manage agricultural and natural resources sustainably, support the health and wealth of Canadian communities and protect the environment (SSC 2016, p. 23).
Central Innovation Hub	To help Departments develop new approaches in a number of areas, including big data.
National Research Council	Developing big data approaches for security agencies and the private sector.
Fintrac	Piloting new ways to use big data to detect money laundering and terrorist financing.

Employment and Social Development Canada	Using big data to assess the labour market, employment insurance programs and improve its overpayment investigations.
Department of Justice	Plans to use predictive analytics in legal services and to determine where to best allocate resources; considering using data analytics to improve individual performance and productivity, and to assess environmental trends to inform policy; discussing using automated tools for risk assessment with cases and in trial and to measure performance.
Canada Revenue Agency	Developing a system to predict benefit eligibility and methods of notifying individuals when they may qualify for benefits; using predictive analytics to identify fraud and non-compliance and to better understand taxpayer decisions; exploring using risk scoring to identify those at highest risk of tax non-compliance; discussion of including behavioural elements and risk indicators from other programs and other parts of the Agency.
Public Safety	Using hot spot mapping to deploy resources at local levels and prevent or disrupt traffic and property crimes.
Royal Canadian Mounted Police	Investigating the use of video capture and related analytics (SSC 2016). The RCMP is also researching how it might use big data analytic techniques such as data mining, sentiment analysis, and social media analysis.
Industry Canada (Innovation, Science and Economic Development Canada)	Using data analytics to improve the way people experience its website and public responses to programs.
Statistics Canada	Exploring how it can make use of unconventional data sources (SC 2016); initiated a pilot project and supported the creation of an Innovation Center that will include a Big Data Lab (SC 2016); investigating new sources of information and data to improve the quality of the Canadian Health Measures Survey data and fill gaps in data (SC 2016)
Citizenship and Immigration Canada (Immigration, Refugees and	Developing a plan to use predictive analytics to assess immigration applications (Keung 2017).

Citizenship Canada)	
Canadian Border Services Agency	Reports using predictive analytics to manage risks but little is known about this program.
Shared Services Canada	Working on infrastructure challenges, and undertaking a scoping project to investigate how it might use predictive analytics to predict and anticipate gaps in cybersecurity and attacks on government services.
Canadian Heritage	Has been investigating social media analytics and developing a digital strategy for data reporting.
Foreign Affairs, Trade and Development Canada (Global Affairs)	Investigating how social media analytics could be used in their forced marriage investigation program; testing how big data could be used in crisis response; a program described as direct diplomacy to gain 'insight into the situational environment' by identifying trends, key issues and influencers; created big data training spaces and programs for staff; developing a tool for Data visualization for fragile and resilient cities; launched an Information and Data Governance Council to prioritize big data initiatives; uses predictive analytics to give advance warning of future events.

* In 2015 the names of several departments were changed. Where this occurred the new name is indicated in brackets in the following table.

Table 7

CHALLENGE OR RISK	DEPARTMENT OR AGENCY (# OF REFERENCES)	DISCURSIVE EXAMPLE
INFRASTRUCTURE AND ACCESS (67)	Canada Revenue Agency	– Difficult for agency to derive maximum value from data as no single entity has responsibility for Agency's holdings
	Canadian Heritage	– Need to ensure big data analytics program integrated with other business-as-usual systems and processes (p.203). – Need for right infrastructure and incentives for innovation
	Global Affairs	– Getting access to datasets within the department, government or externally a challenge
	Interviews	– Government data too siloed to make use of data, need for linkages between data sets (PSD and PSE 2013, PSF 2013, PSB 2013) – There are significant legislative frameworks protecting use of certain types of data that is an issue for linking up data (PSD and PSE 2013, PSN 2015) – There are a lot of competing priorities for resources (PSD and PSE 2013) – The 'challenges are capacity, we're in a period of fiscal restraint and our budgets from an IT perspective do not grow on a regular basis, nor do budgets in program delivery grow, so everybody is under the watchword restraint. And doing new things on that scale requires capacity. It requires financial resources for computing capabilities and software and also the skillsets with teams ready to focus on those. And that involves – in today's environment – having to make some tough decisions as to "Well what will I stop doing if instead I want to work this?" (PSA 2013) – Legislative obligations to ensure data is used properly, and the federal government has to ensure we manage our data assets effectively and in keeping with our legislative obligations (PSA 2013) – In relation to IT the more money that's spent the more politicized, there 'are a lot of entrenched interests in terms of solution providers to federal government and other governments' 'I felt there was a lot of blockage of moving forward with solutions that could help with open data, that could help with making steps towards big data because … other tight top-down control of highly politicized strategic directions and where they're going with technology (PSC 2013). – Entrenched views regarding open source is a frustration. '[A] personal frustration of mine is that there was a lot of long, unnecessary battles to try to move IT and government forward using more mobile, more agile, more flexible programs. Because there's just a lot of entrenchment with proprietary and vendor locked-in solutions that make it hard to do new things like address open data and big data' (PSC 2013) – The cost recovery model prevents sharing of information between departments (PSC 2013) – Lobbying of big tech companies can influence decisions (PSC

		2013) – The Executive Exchange Program can be used in ways that can create bias regarding vendors (PSC 2013) – Internal technical plumbing can be a challenge to making use of big data (PSC 2013) – It can be difficult to know where the data is and acquire it (PSN 2015)
	Justice	– Significant investment will be required in money and human resources needed to avoid falling behind (p.48). – Complex matrix of legislation and practices that surround the use of personal information in public and private sectors in Canada (p. 11, p.36, p.43) – Laws may need to changed to adapt to uses of big data (p.11, p.54) – Implementation of big data a challenge for organizations the size of JUS (p.15, p.231) – 'There is some resistance to providing the necessary data' (p.24). – Expense of building a complete and accurate Big Data store of cases and precedents is a major impediment to predictive analytics in JUS and Canadian legal profession (p.27, p.248). – Determining which jurisdiction responsible for personal information becoming more complicated as information is gathered and/or transferred across legal jurisdictions and co-mingled in Big Data stores or linked with other information sources. Easy to lose track of the origin of the data over time (p.37) – Adequate resources have not been targeted to big data in JUS, need for more reallocation of resources (p.233) – No staff taking responsibility for big data in JUS (p.233) – Not enough big data expertise (p.233) – Lack of access to data (p.235) – Decentralized ownership of data in JUS (p.235) – Lack of formal data standards (p.235) – Data management never been a priority of JUS (p.235)
	Privy Council	– Need to ensure public service has capacity to use big data, requires IT infrastructure (p.3, p.20) – 'Difficult to access relevant databases due to legal, policy and cultural barriers' (p.15, p.10). – Requires re-examination of tools and techniques (p.8) – Creates pressure for real-time interventions (p.15) – Lack of data (p.17)
	Public Safety	– Concern about 'necessary technology, information sharing mechanisms and human resources to use tools' (p.7, p.19) – History of challenges with regard to information sharing (p.16) – Need to improve governance and organization of data and databases (p.20)
	Shared Services Canada	– List of range of data analytics challenges listed in short form (p.4)
	Statistics Canada	– The Social Domain Record Linkage Environment linkages so large and complex that time consuming to add a new file,

		requires big data solution (p.35) – In relation to specific project note need to have access to retailer scanner data (p.42) – Lack of available space on server to process data (p.45) – Concern about inability to keep data suppliers engaged and cooperative (p.45) – Rules governing collection, use and disclosure of confidential business information very strict and pose challenges for realization of the potential of big data (p.154). – IT infrastructure required for big data costly and needs dedicated personnel (p.157)
	Treasury Board	– Real time insights create pressures for real-time interventions that could strain departmental resources (p.15) – Existing legislative and policy frameworks making data sharing difficult (p.19, p.64, p.87, p.128) – Need for necessary IT infrastructure (p.19) – Financial obstacles (p.87) – Improving governance of data and better organizing internal databases (p.19) – Software and technical problems (p.87) – Licensing for the use of big data (p.64) – Some departments are seeking MOUs before sharing data or charging cost recovery (p.87)
PRIVACY AND SECURITY (48)	Canadian Heritage	– 'If done wrong the result will be poor decisions, breaches in data security and privacy, and damage to an organizations reputation' (p.189, p.36)*. – Big data raises issues of privacy and consumer protection, – Limits of anonymization and de-identification (p.192-193, p.3) – Cybersecurity (p.36) – Risk to individual privacy through digital communication, social media and internet of things (p.3) – 'Canadians have a growing sense of unease regarding their ability to control and protect their personal information (p.1). – Safeguards in place to address public concerns about big data (p. 204) – Expected Canadians will become more concerned about privacy as sharing of big data by overnment becomes more commonplace (p.254)
	Finance	– Threat of data breaches (p.80)
	Global Affairs	– Protection of data an issue (p.13)
	Interviews*	– Privacy and governance of data is an issue going forward (PSD and PSE, 2013) – Data security and theft is an issue (PSC 2013) – Concerns about use of personal information and collection of personal information. 'Once you start making decisions about people that affect people's lives based on information that you may not have had the authority to collect, and you certainly didn't make the effort to ensure is accurate, that becomes a secondary problem' (PSJ 2013) – Departments have to do a privacy impact assessment if

		undertaking a program that touches on personal information (PSJ 2015). – Sense that right now the risk is too great and this is influencing people not to be the vanguard in using big data (PSJ 2015)
	Justice	– 'Big Data in government runs contrary to the concepts of personal data privacy' (p.10) – De-identification protections may disappear when data is combined with other sources (p.10, p.194) – Potential for negative public opinion (p.10, p.39, p.45) – Demonstrated ability of hackers to overcome government security, perception personal information is at risk, undermines public confidence in government o holding their data (p.39) – Canadians will be concerns about security and privacy of their data held by government (p.11, p.41, p.234) – Bad news about government surveillance and data breaches my lead negative opinion to go viral (p.11, p.42) – A 'government department that uses social media to try to identify and better understand the needs of Canadians might also be accused of spying on its citizens in order to suppress potential resistance' (p.29). – Difficult to secure consent for secondary use of data, grey area for what information might require explicit or implicit consent (p.40) – 'Achieving full openness while maintaining appropriate controls over data privacy may be mutually exclusive objectives requiring some compromises (p.43).
	Privacy Commissioner	– 'The technologies of collection and analysis that fuel big data are being used in every sector of society and the economy. Online and offline data are combined, which results in a massive increase in the amount of intimate information compiled about individuals. This makes data breaches more likely and more consequential (p.6). – Current privacy frameworks not enough to address developments in health care (p.4) – Gathering broad datasets to catch criminals can sweep up detailed information about people. Need to balance protection of public safety and also civil liberties and privacy of citizens (p.4-5).
	Privy Council	– Need to ensure customization doesn't lead to creepiness (p.15) – Privacy concerns (p. 9) – Need to ensure respect for rule of law and charter rights (p.15) – Anonymization a challenge in context of big data (p.20)
	Public Safety	– Public concerns for security, 'big brother' and 'nanny state' (p.3) – 'Questions of privacy and data increasingly being litigated' (p.3) – Difficulty 'conveying privacy and security risks to users' (p.3) – How to balance privacy and opportunity (p.3, p.19) – Consent and purpose specification (p.6)

		– Security time consuming and challenging. 'Protected B databases to be received from OGDs require burdensome security procedures' (p.87, p. 128) – Need to plan for security and privacy concerns so that data integrity is properly maintained (p.18)
	Shared Services Canada	– Complex security management and integration requirements (p.67)
	Statistics Canada	– 'Perceived lack of consent due to data acquisition may undermine public trust (p.59). – Big data intersects with cyber-security because firms with ever growing databases about clients increases extent of damage caused by security breach than in the past, risk of being hacked as deterrent to big data adoption, hackers can do more damage by using big data analytics (p.156). – Regarding anonymity and re-identification: it can be difficult for an organization to determine how much information stripping is sufficient to meet legislative requirements. Possible to re-identify persons by combining multiple datasets or using advanced software (p.154) – Linking databases could result in detailed portrait of an individual is invasive of privacy (p.155) – Need to ensure customization of services does not lead to 'creepiness' (p.155). – Privacy as major issue identified (p. 1030)
	Treasury Board	– Risk of security breach and being hacked (p.126) – Privacy concerns and data literacy (p.9) – '"big brother" factor' (p.11); creepiness 'Departments must ensure that customization does not lead to "creepiness", where clients wonder how the program knew so much about them (p.14) – Ensuring appropriate respect for rule of law, Charter rights (p.15) – Anonymization a challenge in connection with some big data techniques, such as social media analysis (p.19) – Impossibility of meeting privacy frameworks requirement that personal data no longer required be anonymized or deleted in a big data context (p.64) – Need for 'complex "anonymization" of some databases to protect privacy (p.87) – Linking databases could result in detailed portrait of individuals (p. 124) – Consumers may become captive to providers and unable to extract data, vulnerable to changes in privacy policy, represents 'consumer harm' for competition enforcement (p.124)
DATA QUALITY AND ACCURACY (40)	Canadian Heritage	– Information extracted from big data depends on its quality. Poor quality will lead to poor results (p.190, p.36, p. 79). – 'Potential biases' (p.36, p.79) – Danger of missing input from significant part of the population, lower income groups and older residents may not have access to newer technology (p.192)
	Global Affairs	– Notes bias and potential for measurement errors … 'the

	capture of data points is always governed by the choices of the big data sources and analytical tool's programmers. Note that in many systems 'data' may not be the original observations but rather the products of an inferential procedure of some kind, which can produce additional bias and white noice (p.4). – Big data leaves out non-digital and nondigitizable forms of information (p.4) – Today's data tends to grow faster than their statistical strength, thus heightening the probability of false inferences and amplified noise (p.4) – Big data can be a powerful tool for inferring correlations, it is not for inferring causality (p.4). – Maintaining and enhancing data quality is a challenge (p.18) – 'Big data analysis outside a laboratory context is built on datasets not originally developed for the purposes to which they are being put. Unlike idealized scientific inquiry, where the research question is specified and the research approach is then developed to answer it, big data thus starts from the available data and asks "what can this dataset tell us that might be of interest" (p.8)
Interviews	– Issues with quality of historical data (PSF 2013) – There is so much data out there it is difficult the most difficult part can be choosing what's going to relevant (PSG 2013) – Depending on what tool you use you get different data, so then you get discrepancy between sources (PSG 2013) – Getting data on what people are interested and using across government websites is a challenge (PSA 2013)
Justice	– Gaps in data and duplicates of data (p.24) – 'predicting the outcome of new legal cases is still an imperfect science because of limitations of the current information available for inclusion. Limitations include settled cases not gone to trial aren't available for inclusion in database; courts may not have decided enough similar cases to permit statistical prediction; algorithms relying solely on quantitative feature weights not sensitive to particular context of a problem; lots of data needed to achieve accuracy, manual process needed to represent text cases in appropriate form for machine learning (p.178)
Privy Council	– Social media users not representative of population (p.18) – Difficult to know accuracy of information (P.18) – 'not 100% accurate' (p.8)
Statistics Canada	– Risk that data gathered from social media may be wrong, noisy, irrelevant, inadequate or redundant (p.72) – Noise in the data, data is not clean, data may be ambiguous (p.72) – 'methodological interrogations must be addressed early in any Big Data project or there is a risk that resources are wasted on data that de facto puts into question the credibility of the research…big data is not under the control of the researchers … big data are collected for purposes that are not always congruent with the research questions under

		investigation. Thus the data demand very close scrutiny (p.111) – Problems surrounding selectivity and representativeness are more common when dealing with big data than more traditional sources of data and require special attention (p.51, p.111, p.169) – If third party providing data need to consider the transparency of data collection and analysis. If the provenance of the data is not well understood then the NSO needs to consider whether they can guarantee quality of statistics derived from that data (p.58) – Data from third parties may have time-related problems (p.60) – Depending file structure some data may not be accessible (p.67) – Methodological issues make up most of the challenges, lack of covariates in the datasets, lack of stability in the data … even good science devoid of mischief and misdirection can be ruined by bad data (p.111). – 'The reality remains that not all sources of data are equal. Regardless of the volume, velocity and variety, an evaluation of the data source must precede any analysis and thus … a fourth V called "veracity" is critical to answering what Big Data mean for official statistics. The provenance and quality of the data must be assessed if the research is to be credible' (p.111) – Eligible clients may miss out on benefits if unintentionally excluded by an imperfect model. Ensuring the accuracy of data informing such decisions is essential (p.155). – Lack of covaries, self-selection and self-reporting biases, lack of stability, opportunity for mischief, size issues and selective reporting of results (p. 103).
	Treasury Board	– Social media users not representative of population (p.9, p.16) – Difficulty ascertaining accuracy of information (p.9, p.11) – Need for 'sound information practices' to ensure data is accurate and easy to use (p.15) – Databases 'poorly organized, contain duplicates, typos, etc. Much cleaning required' (p.87)
SKILLS NEEDED (28)	Canada Revenue Agency	– Need for specialized skillsets (p.3) – Difficult to keep up with rapid advancement of business intelligence, may require skill-set not in-house or additional capacity (p.8)
	Culture	– Extracting value maybe difficult and expensive depending on skills and technology an organization has (p.78, p.1). – Shortage of analytical and managerial talent (p.203).
	Finance	– Likely labour shortage in area of big data (p.7)
	Global Affairs	– Scarcity of people to undertake big data analytics, need for internal skills to ensure adequate oversight (p.8) – Need to build financial and human resources to meet the need (p.207) – Need to create new Big data career paths as challenge (p.18)

	Justice	– Staff will need to become experienced with eDiscovery tools (p.20) – Increased need for staff skilled in predictive coding and structured data analytics to analyse complex cross-agency information systems (p.248)
	Privy Council	– Labour intensive (p.8) – Learning curve (p.8) – Lack of staff skills in this area, requires training and collaboration (p.19, p.20)
	Public Safety Statistics Canada	– Need to ensure necessary skills, tools, data (p.7, p.19) – Anticipated skills shortage mean Canada would not be able to fully seize the benefits of big data (p.155) – Applications of big data for policy development are not as intuitively obvious and will involve a steeper learning curve (p.152) – Passively receiving data that is automatically generated as complication (p.154)
	Treasury Board	– Labour intensive and requires learning curve (p.8, p.121) – Staff in corporate management don't have skills in statistical modelling making it more difficult to implement big data analytics (p.18) – Different mix of skills required to use big data and necessitates collaboration (p.19, p.87, p.119) – Need for consultants which means spending (p.87) – Lack of people in Canada with skills needed (p.125)
CULTURE (20)	Canada Revenue Agency	– 'The management and retention of these special individuals requires changes in mind-set and culture' (p.4)
	Canadian Heritage	– 'changing an organization's culture to embrace analytics and adopt a big data strategy' (p.204)
	Justice	– Implementation 'will require a strong change management program, based on the inherent resistance of people to significant change … such a program is likely required in order to achieve widespread adoption of new technologies, and also systems that attempt to measure individual performance more closely' (p. 48). – 'There is cultural resistance to data transparency within JUS. There is a lot of good information that can be shared with the public but there is internal resistance to "put it out there". There is an active resistance to sharing information within JUS. When it comes to data sharing, the default mode at JUS is "secrecy" … Some staff fear that if certain information were freely available, they could be out of a job…. There is a cultural fear of negative exposure within JUS…. There is a culture of risk aversion and a fear of error within JUS'(p.231, p.236).
	Global Affairs	– 'A key challenge in seeking to make use of data to enhance organizational effectiveness is that individuals and organizational entities that have specific knowledge may be reluctant to have the information widely and easily available.

		While such reluctance may be a matter of organisational rigidity or inertia, there are other, very real concerns to take into account such as privacy (including legal restrictions) in regard to the use of some databases. However, even such serious concerns can be mitigated' (p.8)
	Interviews	– 'we have to move to a spot where we (as a government and in relation to making more data open) are more transparent and accountable …. So how do you actually flip that switch, and how do we make that culture change within government to go forward (PSF 2013)
	Privy Council	– Need for 'fearless advice' (p.18) – Avoid perception policy based on popular opinion rather than evidence (p.18) – Need to foster culture of data sharing (p.20)
	Public Safety	– 'to leverage it (big data) effectively you need to create an enabling environment including the right technology, people, and information sharing and data management practices' (p.11)
	Shared Services Canada	– Big data disruptive so need to anticipate system impact (p.12)
	Statistics Canada	– 'as we move to new methods of collection such as web-scraping and the procurement of scanner data, there is a need to clarify the extent to which the Statistics Act, and in particular Section 13, gives us the authority to acquire corporate micro data' (p.115). – Key challenges associated with realizing the potential of big data in the current legal context span public and private sectors (p.153).
	Treasury Board	– Need to ensure 'fearless advice' (p.16) – Avoid perception policy is based on popular opinion rather than evidence (p.16) – Need for 'fostering a culture of data sharing enabling more cross-disciplinary work' (p.19) – Culture of risk aversion preventing data sharing (p.87, p.128) – Requires change of culture to be more open to analytics and data sharing (p.119-120)
CITIZEN AND CONSUMER POWER (12)	Justice	– In list of ethics and values applied to use of big data: 'Full transparency in the application of big data' (p.40) – 'Use power responsibly (when one accumulates information you can exercise power)' (p.40) – It may be more difficult to determine whether information is being used in ways that don't require some form of additional consent or opt-out capability and there may be unintended consequences (p.41) – 'Sharing of information will create complex inter-departmental relationships that will make transparency more difficult (p.248).
	Privacy Commissioner	– Details included in credit scores that make use of 'big data' or algorithms, are largely invisible to consumers and so no

		'meaningful avenue for either identifying harms or holding any entity in the decision-making chain accountable' (p.6) – Lack of transparency of data brokers (p.24) – Government bears responsibility for how big data should be put to use in the public sector and what controls and limits are needed (p.3) – In summary of White House report that refers to uses of data in education, much of it via for profit firms, raises questions about who owns the data and how they can be used (p.4) – Use of big data by government 'increases the potential of the government power to accrue unchecked. Many of the laws governing enforcement were passed before the digital age' (p.6)
	Statistics Canada	– Citizens may not be able to delete data from social media providers, may become captive to original provider who can abuse situation by imposing new fees or changing privacy settings (p.155) – Transparency could suffer if opaque computer models and algorithms are used to inform program decisions (p.155) – Frontline program delivery personal may be unable to explain decisions without help of analyst who build model
DISCRIMINATION (10)	Justice	– 'The potential to profile, target or discriminate against vulnerable people or groups might be possible through matching of open data sources with information gained from other private sources' (p.194). – 'deliver fair and efficient legal operations' (p.40).
	Privacy Commissioner	– Potential of encoding discrimination in automated decisions (p.5) – Lack of transparency and accountability means individuals have 'little recourse to understand or contest the information that has been gathered about them or what the data suggests. This 'creates the potential to discriminate against the most vulnerable segments of society under the guise of neutral algorithms' (p.6). – Discrimination can be intentional and unintentional (p.6)
	Statistics Canada	– 'Big data facilitates profiling, discrimination and automated decision-making, giving rise to concerns about the accuracy of information used for decision-making, accountability for automated decisions, and recourse for erroneous decisions. These concerns are exacerbated by the fact that big data analytics often occur without individuals' knowledge and utilize non-transparent algorithms and metrics (p.154).
	Treasury Board (1)	– The profiling of consumers could lead to segmentation in detailed and subtle ways which could be used to discriminate based on race, gender, religion or other dimensions (p. 124)
	Interviews	– [T]he way I would describe it is that big data is an input to a larger policy-making process that draws on numerous sources of evidence before advice is decided upon. It's also a contributor to better understanding of the way programs are unfolding, in terms of helping make adjustments on the fly. But I can't see anybody relying exclusively on data analytics

		to make any decisions. I mean it's an additional tool in the arsenal, as opposed to kind of wholesale abandonment of the way we did things before and then taking on a data-driven approach to problem solving (PSL and PSM 2015). – A big part of the problem is imagination and knowing where to use the data to augment existing tools to either increase agility or open up new possibilities in terms of government responses, programming or services (PSL 2015). – Imagination depends on context, if you're data rich you're probably less likely to believe in correlations and want to stick to what you know, if you are data poor you are probably more willing to try other things but don't have the knowledge to know how to use data analytics (PSL 2015)
KNOWLEDGE (7)	Canadian Heritage	– 'As we increasingly rely on big data's number to speak for themselves, we risk misunderstanding the results and in turn misallocating important public resources' (p.192).
	Interviews	– Understanding the implications of big data is very uneven (PSB 2013) – Concern about surveillance and related to the Snowden revelations limit how people understand big data and its potential (PSB 2013) – [T]his whole range of policy and ideologically driven motivations that cloud and complicate the scope and visibility even to what they can do with big data or what they can release or what questions they're going to ask of big data' (PSC 2013)
	Statistics Canada	– 'increasingly Big Data are being viewed by developed NSO's as a source of complementary and supplementary information, rather than as a direct replacement for official statistics. While many NSO's are continuing to investigate the possibilities of Big Data, their impact, while certainly important, will not so much transform how we do statistics, but will rather cause our understanding to evolve' (p. 68). – 'there remains much to do before we can definitely say that a place has been found for Big Data at Statistics Canada. We need to define the scope of what types of analysis are acceptable for Official Statistics, which quality assessment framework should be applied to the data, what are the costs and benefits of acquiring and using Big Data and which applications of Big Data technology best helps fulfil Statistics Canada's commitment to produce accurate and timely information that matters to Canadians' (p. 101). – 'As with many National Statistics Organizations (NSO) around the globe, Statistics Canada is yet at the exploratory stage with regards to Big Data. While efforts to increase our knowledge continue, the techniques and technologies surrounding Big Data are generally foreign to most analysts at this institution…. Data sources are … quite unfamiliar with little information surrounding their quality and, indeed, in many cases the data may be best described as completely unstructured. These voids in our current knowledge base are reflected in many NSO's and have lead to a number of international efforts to expand our collective knowledge and

		share our experiences with the goal of developing some common understanding regarding this new paradigm (p.102).
MISSED OPPORTUNITY (6)	Justice	– Large law firms that fail to plan for implementation of new technologies will be at a significant competitive and cost-effective disadvantage (p.150) – Need to be competitive and cost effective (p.46)
	Privy Council	– If only relying on Statistics Canada and other institutional sources of data missing out on new data available and these 'other sources of information will crowd you out' and compete for the attention of politicians trying to anticipate what is happening (p. 14)
	Public Safety	– Reduced capacity in service delivery and efficiency, and Lack of integrated strategy (p.24)
	Statistics Canada	– Low adoption of big data manifestation of broader pattern in Canada of low adoption of technology generally which is detrimental to innovation and productivity (p.147)
	Treasury Board	– Missing opportunity in terms of information and influence (p.13)
ETHICS (6)	Canadian Heritage	– 'The ethical dimension of big data analytics is also becoming increasingly important. With the rule of ethics "just because you can, doesn't mean you should," being raised' (p.80)
	Justice	– Ethical and moral questions about how big data might be used by government or others. 'The more granular the information becomes, the more organizations might be tempted to use the information in negative ways' (p.45). – 'The reasonableness of security arrangements adopted by an organization must be evaluated in light of a risk assessment including a number of factors, such as the sensitivity of the personal information; the foreseeable risks; the likelihood of damage occurring and the resulting harm caused; the medium and format of the storage method, and the cost of putting preventative measures in place' (p. 191). – Ethics and Values that will be applied to the use of big data at JUS detailed (p.40). – 'The ethic of "do no harm" should be applied to how JUS uses data and information' (p.40)
	Statistics Canada	– 'An additional challenge arises from the need for strong values and ethics among data scientists, who have access to massive databases and possess expertise on how to extract information from them. Privacy laws and governance mechanisms can mitigate risks of wrongdoing, but only a strong sense of ethics will help minimize abuse (p.156).
CITIZEN DATA LITERACY (5)	Canadian Heritage	– Ensure the economic benefits to users, organizations and the economy are properly understood (p.204)
	Justice	– People do not have a clear understanding of how much information is being collected (p.234)

	Office of the Privacy Commissioner	–	'consumers most often do not understand how their information is aggregated, sold and used, and the degree to which they are a commodity at each level of this (data) marketplace' (p.5).
		–	Lack of transparency and accountability challenges idea of meaningful consent (p.5)
	Privy Council	–	Data literacy (p.9)
LACK OF LEGISLATIVE AND POLICY FRAMEWORK (5)	Canada Revenue Agency	–	'determining controls to mitigate the risks, needs to be carried out in a BI context (p.13).
	Finance	–	Need for policies around collection, management and use of data (p.9)
	Justice	–	Canadian Charter 'didn't anticipate world where individual's a world where an individual's personal information could be captured and stored in such minute detail, nor the ways in which it might need to be specifically protected' (p.14).
		–	Need for appropriate governance structure so organization has knowledge of what electronic information they have an where it is stored (p.248)
	Public Safety	–	'no clear policy framework has been developed to guide the adoption and implementation of big data analytic tools, in particular the use of social media analytics' (p.10)
ENSURING HUMAN INFLUENCE (2)	Privy Council	–	Need to incorporate human judgement (p.15)
	Treasury Board	–	'Incorporating judgement / human factor in enforcement decisions' (p.15)
REPLACEMENT OF WORKERS (1)	Treasury Board	–	Elimination of jobs through automation of tasks, more likely to replace white collar workers such as legal assistants (p.125)

* Where there are two (or more) page numbers listed in brackets this indicates there were two mentions in the document that were similar. These instances are counted as two mentions.
* No risks or challenges were identified in the Citizenship and Immigration document which is why there is no reference to Citizenship and Immigration in this table.
* Where there were multiple versions of the same document, the risks and challenges identified in the document were only counted as mentions the first time they appear. All versions of each document were analysed, but references to risks and challenges were only 'counted' if they were new and different from what appeared in the first version of the document.
* In order to maintain anonymity interviewees are not attached to any specific Department. The interviews referenced in this table only include those interviews done with public servants.