Open Data Guidebook

A Guide to the Open Data Directive, 2015

TREASURY BOARD SECRETARIAT
Corporate Policy, Agency Governance and Open Government Division
Open Government Office

November 27, 2015

Share this Guidebook

Table of Contents

Share this Guidebook

Share this Guidebook

Purpose of the Open Data Guidebook

The Open Data Guidebook provides Ontario ministries and provincial agencies with information on Ontario's open data processes. It explains the requirements of Ontario's Open Data Directive and provides guidance to ensure consistent approaches to open data practices across government.

The guidebook also provides direction on how to identify data, assess risks, prepare open data for approval and publish datasets. It includes several resources, ideas for open data engagement, templates and answers to frequently asked questions.

Open data in Ontario

Ontario creates, collects and manages datasets covering a wide range of topics. We post many of these datasets as open data online in our Ontario Data Catalogue for public access and use. These datasets are released as open data in accordance with Ontario's Open Data Directive and under Ontario's Open Government Licence - Ontario.

The Open Data Directive requires all data created, collected and/or managed by ministries and provincial agencies to be made public as open data, unless it is exempt for privacy, confidentiality, security, legal or commercially-sensitive reasons.

Ontario Data Catalogue

On November 7, 2012, Ontario launched a public Ontario Data Catalogue.

As part of its commitment to Open Government, the government committed to release new datasets in the Ontario Data Catalogue and directed ministries to develop detailed work plans and metrics to track the release of new datasets.

The Ontario Data Catalogue builds on established best practices and international conventions for making specific types of data open and available to the public through standard formats and a licence with very few restrictions.

Share this Guidebook

Responsibilities

1. **Government of Ontario ministries and agencies** – publish all data that they create, collect and/or manage online on Ontario's <u>Ontario Data Catalogue</u>, unless data cannot be public for privacy, confidentiality, security, legal or commercially-sensitive reasons. Contribute a list of all datasets for which they are accountable for to Ontario's inventory list of data.

2. **Ontario Treasury Board Secretariat (TBS)** – create and maintain an online inventory list of all Government of Ontario data sets, manage Ontario's <u>Ontario Data Catalogue</u>, provide oversight and report on compliance with the Open Data Directive.

How to identify data

Data is defined as facts, figures, and statistics objectively measured according to a standard or scale, such as frequency, volumes or occurrences. Data does not include information like reports or manuals, as explained in the Open Data Directive.

As part of regular government operations, ministries and provincial agencies create, collect and/or manage different types of data.

- Corporate data – workforce characteristics, financial performance, asset characteristics, procurement
- Operational and program data – customer behaviour, program costs, service locations
- Policy and strategy data – market data, economic competitiveness, industry tracking, supply and demand projections
- Evaluation and performance data – program performance data, broader public sector service delivery reports, web traffic statistics

It is important to distinguish data from information. Data is used to produce reports, manuals and information. Wherever possible, reports and other information should be reviewed for proactive release as "open information" under the <u>Open Government Licence - Ontario</u>.

The following questions can help ministries and agencies identify data:

1. Does your ministry have a legislative mandate to produce, collect or share data?

Share this Guidebook

Example

Name: Facility Greenhouse Gas Emissions Summaries

Description: The Ministry of the Environment collects emissions data from regulated facilities for sources listed in Ontario Regulation 452/09.

Link: http://www.ontario.ca/environment-and-energy/2011-greenhouse-gas-ghgemissions-reporting-facility

2. Does your ministry maintain or own the datasets used to inform policy development or program delivery?

Example

Name: Reports on Long-Term Care Homes

Description: The Ministry of Health and Long-Term Care runs its Long-Term Care Homes inspection program to ensure Homes comply with provincially legislated standards. It uses the data that it collects to issue reports and create profiles for each Home.

Link: http://publicreporting.ltchomes.net/en-ca/homeprofile.aspx?Home=C501

3. Does your ministry share statistics, tables, graphs or other visualizations with the public?

Example

Name: Air Quality in Ontario Reports

Description: The Ministry of Environment produces annual reports using data that measures the amount of six common air pollutants as well as the results from the Air Quality Index and Smog Alert programs. The data is visualized using charts and graphs in the Air Quality in Ontario Reports that are published on the ministry's website.

Link: http://www.ene.gov.on.ca/stdprodconsume/groups

4. Does your ministry receive frequent Freedom of Information (FOI) requests or requests from other stakeholders for datasets?

Example

Name: Education Quality and Accountability Office Data Portal

Description: The Education Quality and Accountability Office makes some data available online to researchers. Researchers who require data not contained in the portal can request this through the Ministry of Education's Freedom of Information Office.

Link: http://www.eqao.com/Research/DataPortal.aspx?Lang=E

Share this Guidebook

5. Does your ministry collect or manage data about stakeholders through surveys or consultations?

Example
Name: The Health Care Experience Survey
Description: The Ministry of Health conducts a telephone survey to Ontarians to understand Ontarians' experiences with obtaining health care in the province.
Link: http://www.health.gov.on.ca/en/news/bulletin/2012/hb_20121001_1.aspx

If you answered "yes" to any of these questions, you definitely have data and you need to list the data in your inventory.

Your ministry or provincial agency may also already publish information that contains data, such as:
- tables of numbers (e.g., financial data, spending reports)
- graphs or pie charts (e.g., trends over time, demographic statistics)
- catalogues or lists of items (e.g., recipes, heritage properties)
- information collected and reported at regular intervals (e.g., wildlife population counts, vehicle accident statistics)
- mapped locations with supplementary information (e.g., hospitals, lakes, campsites)
- contact information (e.g., lists of addresses, hunting and fishing licence issuers)
- simple searchable databases (e.g., licensed childcare providers, approved driving schools)

Data that is included in information sources like the above should also be added to the inventory.

How to prioritize data

Ministries and provincial agencies are expected to focus their work efforts and prioritize high-value datasets for release.

There are several factors that can help you decide what data should be a priority for release:
- **Public feedback**
  Data is considered of high interest or value if it is the subject of a high number of

Share this Guidebook

public website searches, related webpage usage (e.g., similar datasets or info), FOI requests or email/correspondence requests. Work with your web team, correspondence unit, call centres, program areas or FOI coordinators to help determine what data the public would like opened.

- **Transparency & accountability**
  Data that increases transparency and accountability, specifically data used to create legislation, regulation, policy or program and service evaluation, or data related to asset management, procurement contracts and audited financial information (e.g., Public Accounts) is considered high-value.

- **Policy evaluation**
  Releasing data that is useful for internal and external stakeholders' evaluation of the delivery of policy, programs or services can be considered high-value. However, usefulness requires timely frequent release in order to support effective evaluation.

- **Government Priority**
  Data that aligns with new or important initiatives is considered high-value.

- **Program efficiency**
  Data that can help improve the delivery of programs and services is considered high-value. To determine if data is of high-value to others, examine the inter-agency or inter-ministerial data requests (e.g., Early Development Instrument).

- **Cost/return on investment**
  Ministries and provincial agencies may want to prioritize data that will cost less to release, such as data that is already published or collected in a machine-readable format and will take less time and resources to prepare as a result. Alternatively, data that has a large return-on-investment, such as data that can reduce FOI requests, create more efficient programs through research, or have positive economic or social impacts, should also be prioritized for release.

Share this Guidebook

Examples of high-value datasets

The following are some examples of high-value datasets based on different topic categories:

Category	Examples
Companies	Company/business register
Crime and Justice	Crime statistics, safety
Earth Observation	Agriculture, Forestry, Fishing, Hunting, Meteorological/weather
Education	List of Schools, Performance of Schools, Digital Skills
Energy and Environment	Pollution Levels, Energy Consumption
Finance and Contracts	Transaction spending, Contracts let, Calls for Tender, Future Tenders, Local Budget, Business Information for Contract Winners: Name, Address, Date Awarded, Amount
Geospatial	Topography , Postcodes, National Maps, Local Maps, Mapped Zones of Designated Permitted Land Use, Property Parcel Boundary Data, Assessed Property Values
Global Development	Aid, Food Security, Extractives, Land
Government Accountability and Democracy	Government Contact Points, Legislation and Statutes, Salaries (pay scales), Hospitality/Gifts,
Health	Prescription Data, Performance Data, Outcomes of Food Safety Inspections
Science and Research	Genome Data, Research and Educational Activity, Experiment Results
Statistics	National Statistics, Census, Wealth, Skills, Government website traffic
Social Mobility and Welfare	Housing, Property Sales, Mortgages and Foreclosures, Health Insurance and Unemployment Benefit

Share this Guidebook

Transport and Infrastructure	Public Transport Timetables, Access Points, Real Time Locations of Transit Services, Broadband Penetration, Public Building Locations, Non-Emergency Service Requests: Graffiti, Noise Complaints, Potholes, Issued Construction Permit Locations, Building Code Inspection Data

Reasons for exemption

Ministries and provincial agencies are required to release all data, unless it is exempt due to privacy, confidential, security, legal or commercially-sensitive reasons. However, exempt data should still be listed in the inventory.

1. Privacy – personal information

A dataset that contains any personal information about an individual must not be disclosed as open data under the Open Government Licence – Ontario. Indirect identifiers can also contain personal information.

Examples of personal information include (but are not limited to):
- race or colour
- national or ethnic origin
- religion
- age
- sex
- sexual orientation
- marital or family status
- medical, psychiatric, psychological, criminal or employment history
- name and contact information (does not apply to individuals acting on a business capacity)
- locations and times of specific individuals.

You might still be able to release data that contains personal information by de-identifying personal info in the data set. This is a process that requires careful analysis and rigorous risk-assessment to avoid re-identification. Your FOI coordinator will be able to help you assess the dataset.

Share this Guidebook

De-identification Process

The TBS Open Government Office is working with industry experts including the Office of the Information and Privacy Commissioner (IPC) and other jurisdictions, to work on a de-identification process.

2. Security

A dataset may pose a security risk if released, by providing information about vulnerable or targeted individuals or organizations. Examples of these datasets may include:

- Information related to policing activities
- Information about vulnerable groups such as prison guards

3. Legal and contractual limitations

A dataset may have legal or contractual agreements that prevent it from being released, such as:

- limitations in data sharing agreements
- third party data – organizations that collect data for a ministry or agency to use, but which the ministry or agency does not have the necessary rights to publish as open data
- commercial licence – data purchased from third parties (e.g. Canada Board of Trade) may have limited rights for distribution
- vendor limitations – external vendors' data services delivered under a contract that forbids the extraction or manipulation of data.
- statute – some data is subject to statutory requirements
- commercial sensitivity – data that may be considered a trade secret

Your legal team will be able to help you assess the data to determine if there are other legal issues.

4. Confidentiality

A dataset may breach confidentiality if its release impairs a government or court's ability to make decisions, such as:

- Court rulings or police investigation
- Budget and policy decisions that may impact financial markets
- Negotiations such as collective bargaining or international trade agreements

Share this Guidebook

Publishing data

There are six key steps to publishing data in the Ontario Data Catalogue.

The table below summarizes the process and identifies who is responsible, what needs to be done and the outcome for each step.

The TBS Open Government Office is working with ministry and agency partners to develop guidance on the publishing process for provincial agencies.

All ministry Open Data must be made available in the Ontario Data Catalogue, which is a one-window central platform for all Open Data published by ministries and is administered by the Treasury Board Secretariat.

Provincial Agencies must publish all Open Data following Open Data requirements.

Open Data Directive, 5.2

Share this Guidebook

WHAT	1. IDENTIFY	2. ASSESS	3. PREPARE	4. APPROVE	5. SUBMIT	6. PUBLISH
WHO	*Ministry lead identifies datasets to be published* **OR** *TBS Open Government (OG) Office contacts ministry (e.g. Top 25 from voting tool)*	*Ministry lead initiates assessment* / *OG Office assesses corporate datasets***	*Ministry lead prepares data for catalogue*	*Ministry lead sends dataset package through approvals*	*Ministry lead submits approved dataset package to TBS OG Office*	*Cabinet Office posts data on catalogue*
HOW	**Ministry lead** contacts OG office to start process. **TBS OG Office** assesses datasets, explains process and identifies participants for next step (assess)	**Ministry** shares data for review with— **Legal branch**: copyright & legal considerations. **Privacy team**: personal & confidential information. **IT cluster**: data quality and accessibility. **Communications Branch**: identifies issues & opportunities. **French Language Services Coordinator**: reviews FLS requirements. **Ministry** determines timelines & reports back to TBS OG Office. ****For corporate datasets**: Horizontal review (TBS + CO) reviews data that is: • Corporate • Financial • Sensitive	**Ministry lead** converts data into open file format with help from IT cluster. **Ministry lead** completes and translates data description form. **Ministry lead** prepares dataset package for approvals. **The dataset package must include**: 1. Briefing Note 2. Reviewed dataset(s) 3. Reviewed dataset(s) description (English and French) 4. Updated HouseBook Note 5. Documentation of approvals (sign-off sheets and/or email confirmations)	The Dataset package is: **Reviewed** by Ministry's: 1. Legal 2. Privacy 3. French Language Services. **Approved** by Ministry's: 1. Communications Branch 2. Manager 3. Director 4. ADM 5. DMO 6. MO (including Legislative Assistant)	**TBS OG Office** receives approved dataset package & confirms requirements. **Cabinet Office Communications** conducts final review of data set title/description. **TBS OG Office** checks in with ministry Communications Branch to ensure their MO is in the loop and has received an updated HouseBook Note. **TBS MO** conducts final check with Ministry MO	**Cabinet Office** updates data inventory. **TBS OG Office** sends links to ministry. **TBS Communications** updates TBS HouseBook Note with total # of datasets. **Ministry Lead** works with Cabinet Office to promote through: • *Social media (tweets)* *Traditional media (links in releases)*

Share this Guidebook

Step 1: Identify data

Identify your data, including those that may be exempt from publication.

Any exempted data should still be listed in the inventory, as well as the reason for the exemption.

Step 2: Assess data

Data must be assessed prior to release for:
- legal, security, confidentiality, privacy and/or commercial-sensitivity issues
- compliance with French Language Service (FLS) requirements, and
- issues management.

Ministries and provincial agencies are responsible for leading the assessment of their respective datasets. For corporate-wide or financial data, the TBS Open Government Office will work with the data custodians to assess the data.

Prior to publishing datasets that contain personal information, the risk assessment process will include tests to determine how that dataset can be used in combination with existing datasets in the catalogue, or other similar datasets that are publically available

Who needs to be involved and how to assess data:

1. **Open Government Lead**
   The ministry Open Government lead should identify their ministry's assessment team, determine timelines, meet each time datasets need to be assessed, and document the results.

2. **Ministry Program/Policy Lead**
   The ministry program or policy lead (e.g., the data custodian or coordinator) should consider and document the dataset's content, any data quality issues, and update the schedule.

   Considerations when assessing data:
   - whether the data can be interpreted by other people
   - whether the data can be used to add value
   - the limitations of the dataset
   - similar data release in other jurisdictions
   - the demand for and importance of a dataset as it may determine a dataset's economic or social impact potential.

Assessment Questions	Considerations
Could data misinterpretation lead to erroneous information being created? Could premature disclosure of the data negatively impact government capacity to provide services and make decisions (government confidentiality)?	Data should not lead to erroneous information being created (flawed/incorrect/mistaken). Data should not undermine regulatory efforts, affect labour relations, and/or compromise budget confidentiality. Consider including a data dictionary and/or releasing information that could provide additional context.

3. Legal Branch

Legal representatives participate to provide advice on the data's legal issues.

Assessment Questions	Considerations
Are the entries in the dataset properly considered data, as it is defined in the Open Data directive?	Only data that meets the definition in the Open Data directive can be released as Open Data under the Open Government Licence - Ontario.
Are there any impediments to releasing the data under the Open Government Licence - Ontario?	All Open Data is released under the terms of the Open Government Licence – Ontario. Therefore, the data must be capable of being released under those terms.
Is the data subject to any statutory restrictions that prevent it from being released as Open Data?	Well known statutory restrictions like the *Freedom of Information and Protection of Privacy Act* (FIPPA) and *Personal Health Information Protection of Privacy Act* (PHIPA) create statutory limits for releasing personal information and personal health information. Other statutes may also create restrictions, such as: • *Vital Statistics Act* – Vital events information. • *Youth Criminal Justice Act* – Information about young persons dealt with under the Act. • *Courts of Justice Act* – Records sealed by a Court.
Is the data subject to any other confidentiality restrictions that prevent it from being released as Open Data?	If the data is confidential to Ontario, it should not be released. Likewise, if Ontario has obligations, such as under contract, to keep the data confidential, then it should not be released. Confidentiality restrictions also exist in the case of

	data for which legal privilege is claimed. Intergovernmental relations or labour relations data may also have confidentiality restrictions.
Do you have the right to release the data publically?	Data is generally not considered property that is owned, such as through copyright. Nevertheless, there may be a contract or a licensing agreement in place governing the use of the data. If this is the case, please ensure the terms allow the data to be published in the catalogue. Work with your ministry's legal branch to determine whether you have the ability to release this data. Consider these questions: • Was the data produced inside by a government employee or outside the government, such as by a contractor? • Is there a valid user/sharing agreement in place that explicitly indicates "do not share"? • If no valid user/sharing agreement exists, have you contacted the third party recently and confirmed a "do not share" agreement?
Are there other restrictions or risks associated with releasing this data?	Could releasing this data create a security risk? Could releasing this data interfere with the administration of justice? Is the data considered classified or sensitive? Would releasing the data violate an internal policy?
Could releasing this data result in potential legal liability?	Downloaders access Open Data under a licence. The Open Government Licence – Ontario does not require developers to indemnify Ontario for anything done by the developer that results in damages to a third party and for which Ontario is sued. The licence also has limited rights for Ontario to terminate the downloader's licence. Downstream users may obtain data from developers, such as through an app. The Open Government Licence – Ontario does not extend to downstream users. Your legal representative should consider the potential for legal liability.

Are there risks associated with future uses?	Under ordinary circumstances, datasets are not to be deleted or removed from public access. Therefore, your legal representative should consider whether the potential legal issues may change in the future and, if so, whether the dataset should be flagged for further review at that time.

4. **Privacy Team**

   Privacy and FOI coordinators assess whether the datasets contain <u>personal and/or confidential</u> information. They also provide background on whether the dataset was included in a FOI request.

   If the data contains personal information, they are required to make recommendations on whether it should be redacted, de-identified or not published.

   The TBS Open Government Office is working with industry experts, IPC and other jurisdictions to develop a de-identification process if your data contains personal information.

Assessment Questions	Considerations
Does the dataset contain any information that can be associated to an individual?	For example, does the data contain email addresses, names, postal codes, etc.? Is the individual acting in their business capacity (e.g., government employee, doctor, registered officiant, etc.)?
Does the dataset require additional privacy measures to be taken? (i.e., aggregation, cell suppression, etc.)	If data contains personal information, ensure that it is removed from the dataset to be submitted for the catalogue. Even if personal information is removed, ensure that individuals cannot be identified if there is a small sample or population count.
Is the release of this dataset in accordance with FIPPA/ MFIPPA (contact your FOI coordinator to confirm)?	Your ministry FOI coordinator can help you to verify if the dataset may contain elements of identifiable personal information and ensure that privacy and security considerations are fully accounted for.
Does the release of the dataset require Privacy Impact Assessment (PIA)?	See: https://intra.sse.gov.on.ca/inetwork/managinginformation/privacy/Pages/pia.aspx

Are personal privacy/public safety threatened?	For example, endanger the life or physical safety of a law enforcement officer or any other person?
	Could the data be used to:
	• expose confidential / personal information?
	• infringe upon personal safety or liberty?
	• violate human rights contrary to the Canadian Charter of Rights and Freedoms?
	• disclose the identity of a confidential source of information related to a law enforcement matter?
	• infringe upon public safety or security?

5. IT Cluster

Ministry IT clusters or ministry analytics branches can assess the data quality and accessibility, and also suggest how to extract the dataset in a machine-readable format.

Assessment Questions	Considerations
Does the data need to be scanned to ensure accuracy of content? Are there data quality issues?	No dataset is expected to be perfect. However, please conduct a review to ensure accuracy to a reasonable extent. Consider these core areas and questions:
	• Accuracy: Does the data reflect real-world objects or a verifiable source? Does the dataset contain errors?
	• Interpretability: Is the data accompanied by the supplementary information and metadata elements necessary to interpret and utilize it?
	• Integrity: Are the relations between entities and attributes consistent? Within tables and between?
	• Timeliness: Is the data out of date? Is the data available at the time needed?
	• Completeness: Is all necessary data present? Are there omissions (gaps) in the data?
	• Access: can the data be obtained from source in a suitable form or medium through which the data can be accessed easily?

Do you anticipate any technical impediments to the linking of the data on the catalogue?	If data is dynamic, how often is the data updated?
	Does it require a read system (API)?
	Where is the data stored (e.g., ministry shared drive, legacy system, cloud, etc.)?
	Are there any planned changes to the collection or storage of these datasets?
	Is the data part of a larger set? Do those sets need to be made available at the same time?
	Do your files meet the required format standards?

6. Communications Branch

Communications can help identify potential issues and opportunities associated with the release of the datasets. They draft House Book Notes, Issues Notes and Q&As to support the release of the dataset.

Assessment Questions	Considerations
Does release of the datasets produce information sensitive to the public opinion?	Think creatively to assess any communications risks that potentially may occur by releasing the data, and prepare an Issues Note if needed.
Can you think of any sensitive or political issues in the media that these datasets will impact?	House Book Notes should be updated whenever a new data set is being released.
Does release of the dataset require an issues management strategy?	

7. French Language Services Coordinator

FLS coordinators will ensure that language requirements under the French Language Services Act are met and will provide advice on identified datasets on French translation needs.

Assessment Questions	Considerations
Has the ministry or program area translated all the necessary information?	Raw data is exempt from translation. However, the data description and any accompanying data dictionaries or documentation must be translated.

Step 3: Prepare Data

The following steps should be taken to prepare your data for approvals and publication:

1. Write a briefing note with background information about the datasets, its relevance, results of the assessment, frequency of updates, issues, mitigation and recommendation.
2. Convert data into open file format with help from IT cluster as applicable. Ensure that specific technical requirements are met and that the data is in machine-readable format.
3. Complete the data description form.
4. Translate the description form, the data dictionary, and other supplementary material for the Ontario Data Catalogue. For greater clarity:
   - Raw data is exempt from translation and is published in the language in which it is collected.
   - The dataset description, metadata, and accompanying data dictionaries or technical documentation must be translated. These will be separate documents that accompany the datasets.
   - The ministry or affected program area is responsible for obtaining and covering the costs of translation.
5. Un-formatted data files that are structured for database consumption already meet accessibility requirements. Formatted tables must pass Excel accessibility checks (e.g., no merged cells, alternative text, properly named columns and sheets, etc.)
6. Update your House Book Note.
7. Prepare your internal approvals documentation.

Step 4: Approve to Publish Data

New datasets have to be approved before they are published. Approvals must be documented, including the name of official and date of review and/or sign off. TBS Open Government Office will require proof of approvals, in the format used by the ministry when you submit your data.

It is important that the dataset is reviewed and approved by the following officials, using existing ministry processes for approvals:

Who needs to review?
- Legal
- Privacy
- French Language Services

Who needs to approve?
- Communications Branch
- Manager
- Director
- Assistant Deputy Minister
- Deputy Minister's Office
- Minister's Office (including Legislative Assistant)

Tip: Identify a key contact in your ADM, DM and Minister's Offices that will help facilitate approvals of open data on an ongoing basis.

What to submit for approvals to your Ministry?
The suggested dataset package must include:
- Briefing Note
- Reviewed dataset(s)
- Draft dataset(s) name/title, description and supporting documents
- Updated Housebook Note
- Documentation of approvals (sign-off sheets and/or email confirmations)

Step 5: Submit Data

Once the dataset is approved, it is ready to be submitted to the TBS Open Government Office. The submission package to TBS Open Government Office must include:
- Data file, link to data files location, or documentation required to access a database directly
- Draft data description form
- Proof of approvals in the format used by the ministry

Before the approved dataset is published online, the TBS Open Government Office will:
1. Assess the data and confirm its compliance with the directive and the guidebook.
2. Liaise with Cabinet Office Communications for a final review of the description form (any revisions will be sent to the ministry for validation).
3. Request French translation once the description form has been validated
4. Check in with ministry Communications Branch to ensure their Minister's Office has received an updated House Book Note.

Step 6: Publish Data

The TBS Open Government Office will work with Cabinet Office Digital to publish the dataset on the Ontario Data Catalogue on Ontario.ca. Cabinet Office will also update the data inventory at this time.

Once published, the TBS Open Government Office will send the relevant links to the ministry and TBS Communications will update the TBS House Book Note with the total number of datasets in the catalogue.

After publication, ministries should work with their Communications Branches and Cabinet Office to:
- promote through social media and traditional media in both English and French (e.g., links in News Releases).
- let their ministry and Ontario Public Service (OPS) colleagues know that they have published a new open dataset
- engage with partners

Tip: Check out the Open Knowledge Foundation's guide for more ideas on what you can do next.

Managing published data

Dataset maintenance
All published datasets must be updated and maintained.

Updates to dataset do not have to go through the entire approval process. However, major structural changes are made (such as the addition or removal of data elements) the dataset must go through the risk assessment process to ensure no risks have been introduced.

> *Ministries and Provincial Agencies must review and update Datasets released as Open Data periodically to ensure accuracy and timeliness.*
>
> *Open Data Directive, 5.2*

To update the dataset, you can:
- Update the data file according to your internal process if the dataset is hosted on your own website and notify the TBS Open Government Office.
- Send the updated file to TBS Open Government Office if the dataset is hosted on the Ontario Data Catalogue on Ontario.ca.

For datasets published as an application (APIs), major changes should be documented in technical documentation hosted on the Ontario Data Catalogue.

TBS will review the published datasets to ensure that previously de-identified data cannot be used in combination with other publicly available data sets, to re-identify individuals. To support this work, TBS will maintain:

1. a list of datasets that contain de-identified data
2. a list of data elements in each dataset

In situations where active maintenance and updates of a dataset are discontinued, or where any published dataset has to be replaced with a newer or modified version, the dataset description should be updated in the catalogue to reflect the "inactive" status, *without* restricting public access to the dataset.

Where custody and control for a dataset is transferred from one program area to another program area, the new program area must update the dataset description in the Ontario Data Catalogue to reflect this.

Dataset removal

Datasets released as open data should not be deleted or removed from public access, except where a dataset is published in error.

If you need to remove a dataset, the following steps must be taken:

1. Document the reason for requesting the removal of the dataset from the catalogue in a briefing note that is approved by your Deputy Minister.
2. Work with your Communications Branch to prepare an issues note and Q&A, to help answer questions about the removal.
3. Submit the briefing note, including approvals by the dataset custodian and the Deputy Minister to TBS Open Government Office.
4. TBS Open Government Office will work with Cabinet Office to remove the dataset from the catalogue.
5. TBS communications will update the House Book Note to reflect the revised number of datasets in the catalogue.
6. TBS will also update the inventory to reflect the reason why the dataset will not be published.

Reporting and compliance

Ministries and provincial agencies must comply with the requirements of the directive. Reviewing and understanding the specific roles and responsibilities in the directive can help maintain compliance.

The ministry and provincial agency is responsible for implementing the directive in their organization. This plan should include accountabilities for ministry and provincial agency staff, and identify contacts for staff to obtain information, advice and direction on potential datasets from the public.

The Treasury Board Secretariat will monitor and report compliance with the directive.

The directive sets out the authority and level of responsibilities to make decisions while maintaining compliance with the principles and the mandatory requirements.

Open Data Directive, 5.4

TBS will conduct an annual review of compliance with the directive through performance measure processes and activities against approved ministry plans. This review may include:

- ministry compliance: percentage of datasets released from the OPS Inventory into the Ontario Data Catalogue
- OPS participation: percentage of eligible ministries publishing data into the Ontario Data Catalogue
- user uptake: increase in web traffic and downloads from the Ontario Data Catalogue
- overall user satisfaction with released data.

TBS will lead the coordination, consolidation and reporting of all compliance and performance measures information when reporting to the public.

Annual Attestations of Compliance from Chair of board governed provincial agencies

As set out in the Agencies and Appointees Directive, Chairs of board governed provincial agencies must confirm their agency's compliance with legislation, directives, and accounting and financial policies in a letter to their minister.

Mandatory requirements of the directive will be built into the annual attestation signed by the CEOs.

The letter submitted by the Chair to the minister will be used by the Deputy Minister to provide assurance of the provincial agency's compliance.

Annual Certificate of Assurance signed by Deputy Ministers
By signing the existing annual Certificate of Assurance Attestation at the end of each fiscal, Deputy Ministers assure that their ministry and all provincial agencies are compliant with all corporate directives and policies.

Open data engagement

Publishing open data is not just about technology, standards, or process. Its key ingredient is engaging with both English and French-speaking data users and the community of data custodians, enthusiasts, and policy makers.

An Open Data Engagement Plan must be developed by ministy and provincial agency to promote and communicate the value and benefits of open data.

> *As part of the Ontario's Open Government Initiative, ministries and Provincial Agencies should develop plans to promote Open Data under the guidance of Treasury Board Secretariat and their specific communication branches. Ministries and Provincial Agencies should communicate the strategic value of Open Data to stakeholders and the public.*
>
> *Open Data Directive, 6*

Here are some tips and guiding questions to help facilitate ministry's and provincial agency's engagement with open data. They are designed to provoke thought and action. To participate in open data events and communities around open data, contact the TBS Open Government Office. .

1. Work within the Ontario Public Service, Broader Public Service and with the public and stakeholders to assist in identifying and prioritizing the release of datasets at the community, regional and provincial levels to support social research and planning.

   Guiding Questions
   - Are your choices about the data you release, how it is structured, and the tools and support provided around it based on community needs and demands?
   - Do you have ways of listening to people's requests for data, and responding with open data?
   - What is the ranking of your datasets on the voting tool?

2. Work with industry stakeholders to raise awareness of the economic potential of Open Data and to prioritize the release of Datasets to support economic activities.

   Guiding Questions
   - Who are the key users of the data?
   - How are they using the data?
   - Do you have a sense of the benefits they are deriving from using Ontario's data?

3. Promote data literacy within educational programs and tools in ministries/provincial agencies under the guidance of TBS.

   Guiding Questions
   - Are you aware of the "data dives" and courses to build skills with data analyses and literacy?
   - Do you provide open data analysis tools and guidance on using the tools?
   - Do you promote, enable and create capacity building to help ministry staff and the community work with open data?

4. Promote the use of government data and digital applications and services with stakeholders.

   Guiding Questions
   - Once published, do you promote your datasets using a variety of external and internal communications channels?
   - Who would be interested in knowing about and having access to your data?
   - Are you aware of any digital applications that have been developed with the types of datasets you manage?

5. Evaluate the potential of sharing the data internally and/or with relevant external organizations in secure environments whenever government data cannot be broadly shared as Open Data.

   Guiding Questions
   - For datasets that cannot be relased as open data, are there opportunities to share with others within the OPS?
   - Are you aware of Statistics Canada Research Data Centres? Are there opportunities to work with them?

6. TBS will coordinate with stakeholders and other levels of government to increase the interoperability and relevance of data. This includes collecting community input, feedback and improving Open Data quality standards.

Guiding Questions
- Can people comment on datasets, or create a structured conversation around data to network with other data users?
- Do you know where the conversations are happening? Do you join the conversations?
- Do you have feedback loops so people can help you improve your datasets?
- Do you collaborate with the community to create new data resources (e.g. derived datasets)?

Compiling the data inventory

Ministries and provincial agencies must contribute their list of datasets to the TBS Open Government Office for the inclusion in the government-wide data inventory. The inventory will include all government and agency datasets that are released, datasets being assessed for release and exempted datasets with explanation for exemption.

The TBS Open Government Office led the development of Ontario's first data inventory in 2013. The TBS Open Government office is reviewing the platform and metadata schema for this second interaction and will share relevant documentation once ready.

The TBS Open Government is also developing a process for provincial agencies to update their inventories. Provincial agencies with websites are encouraged to post their data inventory file along with their published open data on their websites at [agency-url]/open or similar url. Agencies without websites will be able to submit their inventories through their agency coordinators.

The goal of the inventory is to collect the metadata elements required to identity data sources and streamline access using the Data Catalogue Vocabulary Standard (DCAT).

Data Catalogue Metadata information may include the following, but is not limited to:

Purpose	Survey Field	Description
Dataset ID	ID	Number generated to identify and track a unique record
	Data set title	Succinctly describes the content of the dataset

	Data set description	Provides relevant information about the content of the dataset
	Ministry, Division and Branch names	Name the division responsible for generating and updating the content of the data asset or provide authoritative information about it.
Dataset Characteristics	Update Frequency	Indicate how often the data asset is refreshed or updated (e.g, daily, annually, monthly, quarterly, archived, as required).
	Source Data Type	Indicate the structure of the file that your source data is stored in.
	Source Data Format	Indicate the file format of the file where the source data is stored in (e.g., csv, txt, spss, etc.)
	Access	Indicate the level of access currently provided to obtain the data (e.g., branch, division, ministry, government, public, selected third parties)
Open Data Potential	Fee Based Access	Indicate if there are any costs incurred to access and/or use of the data and reason (i.e., contractual agreement, MOU). If there is a cost incurred, provide a rough estimate of what it might be and if this is a one-time cost. Also indicate if fees are for cost recovery purposes.
	Legal Limitations	Indicate if Ontario has the legal right to release this information
	Personal Information	Indicate if the dataset contains personal information
	Sensitive Information	Indicate if data asset contains sensitive information. The dataset can Indicate if data asset contains sensitive information.
	Source URL	Indicate a link to the datasets or information about the dataset
	Additional Comments	Provide any other information about your data asset(s) that would help us to assess its suitability for open data or for listing it in the inventory.

Maintaining data inventory

Once all the data have been listed, the next step is to include all ministry and provincial agency inventories on the government-wide inventory. It is very important for ministries and provincial agencies to keep their inventories up to date – the inventory can be updated at any time.

Ministries and Provincial Agencies are responsible for keeping their Inventory list up-to-date on an ongoing basis.

Open Data Directive, 5.1.6

Procurement and contract data

Publishing procurement and contract data helps make procurement processes more transparent and clarifies government decision-making around procurement. Wherever possible, contract information should be reviewed for proactive release.

The Open Data Directive has requirements for all procurement awards.

Open data as output of contracts

Data collected by third-party suppliers on behalf of the government are required by the directive to be published as open data.

The Request for Bids template posted on the Supply Chain Ontario Tools and Template webpage under "Procurement Documents" includes clauses that will permit your ministry to publish the data collected from vendor(s).

Ministries not using the Request for Bids template are required to work with their procurement and legal teams to develop and include necessary clauses in new contract agreements that will permit government to publish data collected.

The Government of Ontario will obtain the right to publish procurement contract data as Open Data. Procurement contract data such as the winning bid for every contract awarded (e.g. vendor name, financial payment information) should be included and published in a timely manner, unless excluded from being made available as Open Data. Vendors shall agree that financial data of contracts are not considered commercially sensitive and may be released.

The Government of Ontario will obtain the right to publish as Open Data original data created or collected as an output of contracts with suppliers, unless the data should be excluded from being made available as Open Data.

Open Data Directive, 5.3.1

In support of open data, the Ontario Government publishes various procurement datasets on Ontario's Open Government Catalogue including but not limited to:

- enterprise-wide VOR arrangements planned for the next three years by Supply Chain Ontario (SCO)
- enterprise-wide VOR arrangements that are managed by SCO

On an annual basis, ministries and provincial agencies are requested to release their Consulting Services and Agreement of Internal Trade procurement reports to SCO for consolidation and publication on Ontario's Data Catalogue.

In support of the open format, ministries looking to acquire new data management platforms, applications and solutions may use the <u>Business Analytics, Data Integration and Quality Tools and Services (BADIQ) Vendor of Record (VOR) Arrangement</u> or its successor, made available through SCO.

Ministries must consult their Open Government Lead, Privacy Coordinator and legal team when considering a potential exemption to disclosing a procurement record.

I&IT requirements

The directive has implications for both I&IT solutions and processes, including procurement and internal builds:

- new data management platforms, applications and solutions must include open data requirements while older I&IT solutions should be updated when possible, to add this function to provide data online in an open format.
- Governance of open data requirements must be built into the I&IT Gating and Architecture Processes
- all I&IT solutions require proactive consideration of privacy protection throughout the lifecyle of the data. For more details, see the Corporate Policy on the Protection of Personal Information.

The TBS Open Government Office is working with experts on the definition of data management platform.

I&IT gating and architecture

Going forward, I&IT project gating and architecture processes will include requirements to ensure new I&IT projects include requirements that enable the publishing of open data.These requirements will be incorporated into the enterprise architecture review requirements, for the relevant projects.

I&IT procurements and solution development

Any new I&IT data management platforms/solutions/applications, whether a new build or the purchase of an off-the-shelf product, must be designed to allow for the extraction of data, making it easier to publish and regularly maintain data in the <u>Ontario Data Catalogue</u>.

Program areas should work with their I&IT cluster and/or procurement advisors to:
- ensure that the extraction of data in an open format through automatic means, whether through a recurring task or an application programming interfaces (API), is included as part of the requirements gathering process.
- include the requirements and necessary clauses in all RFPs/RFSs and contracts/service agreements.

Best practices for extracting data include:
1. Providing a linked, rich API over HTTPS in an open format through an API, and publishing the URL to the API in the <u>Ontario Data Catalogue</u>. The URL should use https to ensure the privacy of the user.
2. Providing a real-time or snapshot view of a particular dataset at a URL.
3. A periodic, automatic report that produces an open format suitable for upload to the <u>Ontario Data Catalogue</u>.

Future considerations

Internal data portal

TBS is exploring options for an internal platform and a toolset to manage, collect and display information about datasets available within the OPS.

The platform pulls information from various sources (e.g., Drupla and WordPress sites) across the OPS and will also integrate with the public <u>Ontario Data Catalogue</u>.

An Internal data portal will:
- Facilitate data sharing within the OPS and enhance the sustainability of open data by integrating it into regular operations, and encourage the use of data, while ensuring that applicable privacy, security, legal, confidentiality, and commercially sensitivity requirements are followed.

- Support the implementation of the Open Data Directive – it will contain the OPS data inventory of all data (open by default or not).
- Help automate key open data processes – including the risk assessment, approvals and publication.
- Enable tools and link to the larger IT infrastructure.

A proof of concept will be ready in spring 2016 and TBS Open Government Office will work with ministry Open Government Leads to migrate inventories into the portal.

Resources and templates

Data Description Form

Fill in the English portion of this template with the details of your dataset and submit to TBS Open Government Office for review. A review with Cabinet Office communications will take place with TBS Open Government Office. Once finalized, you will be responsible for French translation and the information will be loaded into the Ontario.ca page for your dataset.

Page Metadata	
Topic (global 13)	e.g., Environment and Energy
Sub-topic	For CO Comms use only
Related datasets	List additional datasets on OneSite by title/URL that would make a set
External tag	For CO Comms use only
Owner/Publisher (ministry)	Ministry of
Dataset ID #	If it was listed in the OPS data inventory
Contact us	Secondary contact:
About this dataset	
Date created	MM/DD/YYYY (optional)
Date published	In Ontario.ca catalogue (system default)
Date modified	most recent update of dataset if different from publication date
Date Range	(single date for data snapshot) MM/DD/YYYY to MM/DD/YYYY
Update frequency	Monthly/Quarterly/Annually/As-Required/other/etc.
Geographical coverage	Ontario/Canada
Technical documentation (Eng)	Metadata Record https://www.javacoeapp.lrc.gov.on.ca/geonetwork/srv/en/metadata.show?uuid=2937e848-c54e-4b34-abde-20a1ca9dd0ab&currTab=simple
Technical documentation (Fr)	Registre de métadonnées https://www.javacoeapp.lrc.gov.on.ca/geonetwork/srv/fr/metadata.show?uuid=2937e848-c54e-4b34-abde-20a1ca9dd0ab&currTab=simple

English	
Page Title	**Webpage title goes here (Eng)**
Lead (limit 200 characters)	Brief overview statement about dataset(s) provided on this page.
Body (should be under 100 words - must be different from lead)	Describe dataset, including any essential background information.
Licence Type	Open Government Licence - Ontario
Licence URL	http://www.ontario.ca/government/open-government-licence-ontario
Filename(s)	
URL (for linked datasets)	
File types	e.g., SHP, XLS
Additional instructions	Put any additional notes or instructions here

French	
Page Title	**Translation**
Lead	Translation
Body	Translation
Licence Type	Licence du gouvernement ouvert – Ontario
Licence URL	http://www.ontario.ca/fr/gouvernement/licence-du-gouvernement-ouvert-ontario
Filename(s)	
URL (for linked datasets)	
File types	
Additional instructions	Add translation of English version and any additional notes or instructions specific to French version.

Open data file formats

Open datasets are formatted as non-proprietary, open files to promote broad use and ensure compatibility when working with other sources.

These tables below describe common data file formats. Not all file formats can qualify as open data. Data formats must be:

- **Platform-independent**
  The file should not be tied to a particular computer vendor.
  - Example: A numbers spreadsheet is only viewable on Apple devices and should be avoided in favour of CSV, which is viewable on any platform.

- **Machine-readable**
  The file formal should be suitable for machine processing.
  - Example: A table in PDF format is intended for a human viewer, not a machine. The underlying table should instead be published in a suitable format for tabular data, such as CSV.

- **Vendor-neutral**
  The file format should not require proprietary software to view.
  - Example: A Microsoft Access MDB file requires Microsoft Access to view. Data should be released as CSV or open database formats such as MySQL SQL files.

- **Standardized** (where possible)
  Certain data types have actual or de facto standards adopted by their communities. Data owners should strive to use these standards where available.
  - Example: Transit data is almost universally released in GTFS. This is preferable to custom formats as it will allow easier integration of Ontario data with existing tools.

Acceptable Open Formats lists file formats that meet the requirements above.

FILE FORMAT	DESCRIPTION
XML	Extensible Markup Language – a set of rules for encoding documents electronically. Well suited for consumption by automated programs capable of handling raw XML (extensible markup language) files. • Simple, general textual data format • Easy to parse via a programming language o Hundreds of XML-based languages have been developed, which means broad range usage over multiple programming languages • Easy to use over the internet
CSV/TXT	Comma-separated values – a file that is used for the digital storage of data structured in a table of lists form. Use these formats for easy access to the data. Text/CSV (comma separated values) files can be opened by most desktop spreadsheet applications (e.g., MS Excel), but are also machine readable. • CSVs are widely supported • Stores tabular data (physical records in plain text formatting) • Common on all computer platforms
KML	Keyhole Markup Language – an XML-based language schema for expressing geographic annotation and visualization. • Used to display geospatial data in Google Earth (originally named Keyhole Earth Viewer), Google Maps and similar applications. KML (keyhole markup language) is extended from XML and shares some of the same structure as GML (geography markup language).
KMZ	Zipped KML files.
GML	Geography Markup Language – the XML grammar defined by the Open Geospatial Consortium (OGC) to express geographical features.
SHAPEFILE (SHP, SHX, DBF, PRJ, SBN)	Used for consumption by shapefile-compatible mapping applications. Most datasets in shapefile format are updated on a monthly or quarterly basis as they are not "operational" in nature. • Popular geospatial vector data format for geographic information systems (GIS) • Stores geometric location, geometrical data types of points, lines and polygons and associated attribute information • Lacks capacity to store topological information
GEOJSON	GeoJSON is a format for encoding a variety of geographic data structures.
GEOTIFF	GeoTIFF is a public domain metadata standard that allows geo-referencing information to be embedded within a TIFF file.
RDF	RDF (resource description framework) is a standard model for data interchange on the web. It is based on XML and URI (uniform resource identifier).
ODF	Open Document Format – XML-bases document format.
OOXML	Office Open Extensive Markup Language – an ISO/IEC standardized file format.
JSON	JavaScript Object Notation is a lightweight data-interchange format.

File <u>formats that are not encouraged</u> are generally usable by the public, but don't meet all of the requirements of Open Formats.

FILE FORMAT	DESCRIPTION
XLS	XLS is a file format used with Microsoft Excel. Main spreadsheet format which holds data in charts, worksheets and macros. • Data can be transferred into a CSV/TXT format.
Microsoft Access file types (MDB, MDA, ACCDB)	Microsoft Office Access is a pseudo-database management system that can combine data from other file formats such as XLS, SharePoint lists, text, XML, etc. to compile in a single database. • The end result of this database file should be converted to a CSV/TXT or another machine-readable format.
ZIP	Zip files are compressed files or folders. Zip files are not encouraged within an Open Data setting, unless necessary (a skin, blog theme) since files included are uncontrolled types. • Supported by all operating systems Windows 98 and later, and Mac OX X 10.3 and later.
Daisy	Daisy is a Java/XML open-source content management system and is in use at major corporations and governments for intranet knowledge bases, project documentation and management of content-rich websites. This content is stored in so-called Daisy documents. • Uses a WYSIWYG (what you see is what you get) editing environment, and is a front-end application. This is not a back-end machine-readable format.
HTML/HTM	HTML (hypertext markup language) is the predominant markup language for web pages. Alone, it is not optimized for machine-readability, but only the output is meant for human readability. • Although some HTML files (depending on the site's usage) can be converted to XML or XHTML to be machine-readable, this does not mean *all* HTML files can be converted. For this difficulty in defining and determining (by the average user) which files are good for conversion and which ones are not, HTML should not be an acceptable file format.

File formats that should not be used.

FILE FORMAT	DESCRIPTION
DOC	DOC files are files and documents that are created by word processing software. Data contained in this type of file should be analyzed to determine which file format to save it in for inclusion within the data catalogue.
Media files (JPG, PNG, GIF, BMP, MP3, AVI, etc.)	Media files should not be included into the database since it does not provide actual source data. Even if the JPG may be a chart or graph, or if the MP3 is of a recent podcast that discusses data –these are not good formats because data cannot be extracted at all.
EXE	An executable file is essentially a full program. This file type should not be released at all.
JS, PHP, ASP, ASPX, PLS, VBS	These files are scripting languages that are used to create dynamic content or functions on web pages. This type of file does not carry any type of actual raw data and should not be included in an open data database.

Sample Briefing Note

Developing a briefing note for ministries is a component of publishing data. The briefing note must be included in the dataset package for approval.

The briefing note should include the following information:

- **Purpose:**
  Request approval to publish datasets identified below into the Ontario Data Catalogue on www.ontario.ca/opendata.
  <List the datasets>

- **Current Status:**
  As part of its commitment to open data, the Ontario government committed to release new datasets in the Ontario Data Catalogue and directed ministries to develop work plans and release new datasets.

  The <Ministry, Division, Branch> determined that the following datasets can be added to the Ontario Data Catalogue.

  <Insert a brief descriptive overview of the datasets>

  <Insert any important considerations/information with regard to risk assessments> Considerations for risk assessments:

    o Have you identified any copyright and legal risks?
    o Are there any personal or confidential information included in the dataset?
    o Is the data high-quality and accessible? Are there communications issues or opportunities?
    o Does the dataset follow French Language Service requirements?

- **Background:**
  <Insert any relevant background, tailoring to your audience. Information on the open data directive, open government action plans, number of datasets etc. can be found on the Open Government website.>

- **Recommendation:**
  It is recommended that the datasets described in the attached Dataset Description Form be published in the Ontario Data Catalogue on Ontario.ca.

- **Attachments:**
    1. Reviewed dataset(s)
    2. Reviewed dataset(s) description (English and French)
    3. Updated HouseBook Note /

- **Approved By:**

Frequently Asked Questions

What is the Open Data Directive?
The Open Data Directive outlines key principles and requirements for publishing open data, and will apply to data created and managed by Ontario ministries and provincial agencies.

The directive aims to maximize access to government data by requiring all data to be made public, unless it is exempted for privacy, confidentiality, security, legal or commercially sensitivity reasons.

The directive will also require a public inventory of all OPS and provincial agency data assets, and require future IT investments and procurement contracts to consider open data requirements.

How does the directive fit into the Engagement Team's recommendations on open data?
The province appointed an Open Government Engagement Team made up of public engagement and digital experts. This team travelled across the province, used a variety of social media tools and traditional media to gather feedback from the public, and spark new conversations about how government can engage, innovate and collaborate.

One of the Engagement Team's recommendations was to implement an Open by Default Data Policy. The Open Data Directive is the government's response to this recommendation.

Who will the directive apply to?
The Open Data Directive will apply to Ontario ministries and provincial agencies.

The directive will not apply to broader public sector organizations. We continue to work with the broader public sector to improve access to their data by sharing best practices.

Will staff be trained on open data requirements?
A wide range of Open Government tools, resources and training – including courses on open data – will be offered to staff to help build skillsets and capacity in this area.

What is the definition of data?

Data is defined in the directive as facts, figures and statistics objectively measured according to a standard or scale, such as frequency, volumes or occurrences, but does not include information.

What does the Directive apply to?

The Open Data Directive applies to all data created, collected, or managed by ministries and provincial agencies.

What is open data and what are the benefits?

Open data is a concept in which government data is made available to the public, where appropriate, so that the public and businesses can repurpose the data to develop new ideas, services and applications. For data to be "open data" it needs to be made available at no cost, in a digital format and under a flexible use license that carries few conditions.

The public and businesses want to access and use government data in order to make better decisions, develop services and applications, and create new knowledge.

There are many benefits of launching an open data program, including:
- Stimulating Economic Growth: businesses want government data for developing applications and services for commercialization, or to understand more about the province, in order to grow their business. Sharing our data in an organized way might even attract certain foreign investors to start their businesses in Ontario.
- Improving government service delivery: when deciding to open up data, government must become more organized in the way it manages data. This improved organization of information reduces search time and facilitates timely decision making. When data is released, citizens and businesses can develop technology solutions and derive insights that can help improve government service delivery.
- Improving public accountability and transparency: sharing data as open data can lead to fewer Freedom of Information requests, lower costs and reduced administrative time. Open data makes it easier for the public to understand what government is doing and accomplishing. This increased transparency of government actions leads to greater citizen engagement and confidence in government.

Do I have to publish the data for every dataset listed in the inventory?
Some datasets cannot be published due to privacy, security, confidentiality, legal or commercial sensitivity reasons. However, these datasets will still be listed in the inventory along with the explanation of why they cannot be published.

Why might a dataset not be published as open data?
The directive specifically exempts data for the following reasons:
- data is subject to statutory confidentiality requirements
- data should not be disclosed for legal, security, confidentiality, privacy or commercial sensitivity reasons.

For assistance, please contact your Freedom of Information and Privacy Unit, legal team or the Open Government Office, TBS.

What happens when a complaint is received about the data being published?
It depends on what the complaint is. If it's about the accuracy of the data or the fact that it is published at all, TBS Open Government Office will refer the complaint to the respective ministry/ provincial agency.

Other non-specific, general complaints will be handled by the TBS Open Government Office.

What do I need to do if I want to see unpublished datasets?
Ministries are responsible for prioritizing and categorizing datasets for publication.

Ministries are required to be "open by default", so feedback and requests from the public should be used to prioritize release of open data.

Ministries should anticipate that members of the public could submit Freedom of Information requests for datasets not released as open data and be prepared to release the data unless it falls under one of the exclusions or exemptions in FIPPA.

When will the inventory process take place?
The TBS Open Government Office is currently working on the inventory process and platform. We expect to start sharing the metadata schema and offer training in winter

2016, and start working with ministries compiling the inventory spring 2016. Updates and maintenance will be ongoing with a yearly quality assurance review led by the TBS Open Government Office.

Do I have to list every dataset in the inventory?
Yes, ministries and agencies must list all datasets in the inventory. The only exception is that ministries may exclude datasets where the:
- data is subject to statutory confidentiality requirements
- data should not be *disclosed* for legal, security, confidentiality, privacy or commercial sensitivity reasons. Data that cannot be published due to these reasons but may still be disclosed must be included in the inventory.

What data is excluded from the directive? Does the directive allow ministries and provincial agencies to opt out through exemptions?
The directive requires all government data to be made public, unless there is privacy, confidentiality, security, legal or commercially sensitive reasons not to release it. This means that Ontario will not publish data that:
- contains personal or confidential information
- has legal, copyright or security restrictions, or
- compromises public safety.

We know releasing government and provincial agency data is the right thing to do. It's also the smart thing. We will work with our partners to ensure they understand the principles and benefits of open data, and follow their obligations through the directive.

Does all data also have to be available in French?
The Dataset Description Form must be provided in both English and French given that the name/title and description in the public-facing data inventory must be available in French.

In addition, the ministry and provincial agency responsible for the data should consult with their French Language Services Coordinator in order to review datasets and discuss French translation needs. Although most raw datasets and numerical data do not require translation, the title, description, metadata, data dictionaries and accompanying technical documentation must be translated in order to comply with the posting requirements.

Can you give me some examples of high-value data sets that you have opened? What do you mean by high-value?

Ontario's datasets cover several subject matters, from transportation to the environment, and include a number of high-value data sets. High-value data sets are deemed to be those that are:

- frequently requested by the public via Freedom of Information requests, public surveys, phone, mail inquiries, and media
- collected or produced in support of government priorities
- legislative mandate to produce, collect or share data, or
- identified as such by reputable international organizations, such as the Open Government Partnership or the Open Knowledge Foundation.

Some of Ontario's high-value datasets include:

- University and College enrollment
- Personal income tax rates and credits
- Traffic volume on provincial highways
- Employment Service Program information and outcomes
- Energy Use and Greenhouse Gas Emissions for the Broader Public Sector

What formats do you use to publish these datasets?

Ontario's data sets are made available in machine-readable formats (formats that computers and other types of technology can easily process). Formats include

- XML (extensible markup language)
- CSV (comma-separated values)
- TXT (text) files

Depending on the data, other formats may be available. Our datasets are also published under our Open Government Licence - Ontario which allows data users to easily compare or "mash" Ontario data with data from participating jurisdictions like the federal government, Alberta, British Columbia and the City of Toronto (who have similar Open Government Licence - Ontario terms).

When is the effective date of the directive?

The Open Data Directive takes effect on April 1, 2016.

My ministry/agency is worried about resource implications to support the implementation of the directive.

The Open Government Office and ministry leads will support ministries in implementing the directive and providing assistance as required. Ministries and provincial agencies are expected to:

- prioritize the release of datasets that address government priorities or have been requested by the public
- optimize resources to focus on high demand datasets
- develop business cases in instances where additional funding is required

What do I need to know about open data requirements when planning an I&IT project?

If your I&IT solution incorporates data management, one of the requirements must be that it can extract the datasets into an Open Format.

The gold standard as identified in the Directive is to provide the data online through an API, and publishing the URL to the API in the Ontario Data Catalogue.

Other approaches to satisfy this requirement could include a button that an admin clicks to generate a machine-readable report, or having the system periodically email a report to the program area.

Projects should ensure these requirements are built into their Requests for Proposals (RFPs) and contracts with vendors; or are available from off-the-shelf solutions. Submissions to the Architecture Review Board (ARB), cluster governance, as well as TB/MBC and ITPAC submissions should highlight how the requirements of the Open Data Directive will be met.

What if I procure a solution instead of building it?

Procured data management platforms are also required to support open data requirements.

Which I&IT solutions are subject to the Open Data Directive? What is meant by data management solution vs. all IT solutions?

In general, transactional systems include data therefore should have open data export functionality built in. Keep in mind that most transactional systems will include direct or indirect Personally Identifiable Information (PII) which will not be eligible for publishing on the Ontario Data Catalogue. Where possible, such systems should be designed in a way to facilitate publishing those portions of the data that does not contain PII. For

example, a database of room reservations might publish the list of rooms but not the reservations themselves.

When it comes to computer logs, it is not the intention of the directive to include all computer logs generated by all systems in the definition of data. However, if the specific character of a log meets the definition of data, it should be included in the inventory. For instance, it doesn't make sense to publish the system log of every computer in the organization. On the other hand, website analytics data such as page views is a dataset and is used to inform communication planning, and should be published.

How do I ensure my ministry/provincial agency is in compliance with the directive?

Ensure that your ministry/agency follows the mandatory requirements of the directive and the review and approval steps outlined in the *Publishing Data* Section.

TBS will play an oversight role by conducting an annual review of compliance with the mandatory requirements of the directive. Ministries will leverage the certificate of assurance process that assures ministries are in compliance with all corporate directives and policies while provincial agencies require attestations of compliance from board governed Chair/CEO.

How do I know if my organization is a provincial agency required to follow the Directive?

If you are a provincial agency, you are required to comply with the Directive. The Agencies and Appointees Directive can help you determine if you are a provincial agency. It defines the characteristics of a provincial agency as follows:
- is established by government through a constituting instrument (under or by statue, Order in Council or regulation);
- is accountable to a minister for fulfilling its legislative obligations, the management of resources it uses, and its standards for any services it provides;
- The majority of its appointments are made by the government;
- Is not organizationally part of a ministry but is part of the Government of Ontario;
- Has authority and responsibility, granted by the government, to perform an ongoing public function or service that involves adjudicative or regulatory decision-making, operational activity, or an advisory function.

You can also check the provincial agencies list for your organization.

What happens if I publish my datasets in error?
There may be exceptional circumstances that require datasets to be removed from the catalogue. The directive states that the approval of the Deputy Head is required for the removal of ministry datasets from the catalogue. Once approval is obtained please contact TBS Open Government Office.

How do I find out who my Open Government Lead is in my ministry?
Each ministry has appointed an Open Government Lead. You can find out who your lead is on the Open Government Intranet site.

Who can I contact if I have more questions about the directive?
The Open Government Office, TBS is happy to provide assistance. You can email opendata@ontario.ca or refer to our contact list on the Open Government site.