



# Open Data Master Plan 2018 - 2022

**JANUARY 2018** 

# **Table of Contents**

Foreword 4	
How the plan was created 5	
Who this plan is for 6	
Why Open Data matters 7	
How we will get there 8 - 9	
Where we want to be 10 - 11	
What this plan is for 12 - 13	
How we will get there 14 - 15	
Theme 1. Foundation 16	
1a. Update publication pipeline 17	
1b. Prioritize data releases 18	
1c. Emphasize quality over quantity 19	
Theme 2. Integration 20	

2a. Establish program governance	21
2b. Allocate resources for growth	22
2c. Incentivize teams to publish data	23
2d. Change cultural perspective	24
Theme 3. Connection	25
<i>3a. Facilitate market opportunities</i>	26
3b. Engage local talent	27
3c. Cultivate community partnerships	28
Theme 4. Activation	29
4a. Enable effective use of data	30
4b. Increase awareness of Open Data	31
When the plan will happen 32	- 33
How to get engaged	34

### Foreword

The City of Toronto envisions a future where we enable anyone, anywhere, to improve life in Toronto with open data. We are focusing on Toronto's core competencies around innovation, diversity, technology and talent to encourage and stimulate the potential of open data and become world leaders in the Open Government movement.

Co-developed with a diverse set of stakeholders and with input from vibrant and engaged partnerships, the plan enhances transparency and facilitates improved public service delivery through participation with citizens, businesses, community groups, not-for-profit and academic organizations, other governments and beyond. The principles of the plan are to focus on datasets to deliver meaningful civic solutions, improve City efficiency and remove barriers to open data.

This Open Data Master Plan is a significant milestone and outlines a framework to guide the next phases of open data in our city. It will help guide us in our journey towards excellence and supports our goal of using open data to improve service delivery and increase communication and data sharing between City divisions. I would like to acknowledge the focused and dedicated effort of members of the public, the open data community, and City staff in the development of this plan and welcome you to join us in the journey ahead.



**ROB MEIKLE** Chief Information Officer

# How the plan was created



5

# Who this plan is for



Toronto's vision for open data

we plan to unlock its potential

leveraging open data

## Why open data matters

When government data is made open to the public, new ideas and perspectives unlock exponential potential for it to be re-used, analyzed, and correlated to help improve the City's delivery of public services, engage with citizens in government decision making, and innovate our approaches to civic problem solving.

A stable foundation and substantial investment in open data is a key contributor to the success of strategic initiatives like Smart Cities, Civic Innovation, and Open Government.

## Where we are

TIMELINE





#### LESSONS LEARNED

- Focus on releasing datasets of high quality and value rather than quantity of datasets
- The open data team requires varied competencies, technical and non-technical, to be effective
- Publication needs to be automated and streamlined to meet growing demand for open data
- Simply releasing open data does not make it accessible to all groups addressing civic issues
- Community groups are critical partners for capturing the value of open data and improving the program
- Data releases must be prioritized to balance resources needed with user value created
- Open data's potential to address civic issues is key to estimating its value
- Open by Default needs to be a formal requirement and holistically embraced by senior leadership

We envision a future where we enable anyone, anywhere, to improve life in Toronto with open data

# Where we want to be

We believe Toronto's core competencies around innovation, diversity, technology, and talent present an opportunity to activate the potential of open data and become world leaders in the Open Government movement.

It is important to acknowledge two points given this ambitious goal. First, given municipal government is but a single component in the vast and complex Toronto landscape, we simply cannot realize this vision on our own without ongoing community support. Our aim is to take ownership of the areas we influence, and facilitate engagement where we can maximize benefit.

Second, in order to be successful the plan must evolve and adapt to a dynamic world; indeed, the pace of innovation has never been more rapid. For instance, the Toronto-Waterloo Corridor's transformation into an international technology hub, complexities in data privacy legislation, and new approaches to data usage agreements with third parties are just a few examples of changing factors that may influence open data.



The primary goal of the Master Plan is to provide a strategic framework and a roadmap to advance our vision for open data until 2022

# What this plan is for



Make open data an official program and adopt the International Open Data Charter. Identify the right resources and infrastructure to effectively grow and advance the Open Data program. Increase the quality, speed and efficiency of open data production. Demonstrate the value created through open data by showcasing successes and insights. Lead by empowering others through effective governance, open source development, and coproduction. Build a foundation for enhancing the ability to address civic issues through increased collaboration and partnership.

# How we will get there

#### **GUIDING PRINCIPLES**

The International Open Data Charter (ODC) is the product of an international collaboration developed to help governments maximize the potential of open data. The ODC sets priorities and benchmarks open data across the world. The City of Toronto's Open Data Master Plan is anchored in the international principles defined by the ODC. This allows Toronto to measure itself on an international level and holds our city to the highest standard. The ODC's six principles were used to guide the multi-stakeholder engagement and ultimately, the master plan.

	OPEN BY DEFAULT	Free access to, and use of, open data is of significant value to society and the economy; hence, it should be open by default unless it compromises personal data or privacy rights
	TIMELY AND COMPREHENSIVE	To be of value, open data should be prioritized given the needs from users and the resources needed for publication, and it must be comprehensive, accurate, and of high quality
ę	ACCESSIBLE AND USABLE	Open data should be easily discoverable, widely accessible, and made available without bureaucratic or administrative barriers that would deter people from accessing it
	COMPARABLE AND INTEROPERABLE	To be most effective and useful, open data should be in structured and standardized formats that enable comparison within and between sectors, locations, and over time
	IMPROVED GOVERNANCE AND CITIZEN ENGAGEMENT	Open data strengthens trust in government by enhancing transparency around decision-making, increasing visibility into how society's needs are being met, and empowering civic participation
	INCLUSIVE DEVELOPMENT AND INNOVATION	Governments must aim to enable everyone to understand and use open data effectively so benefits generated from it are not limited only to those with specific tools and resources

#### INTERNATIONAL OPEN DATA CHARTER ALIGNMENT

FOUNDATION			INTEGRATION			CONNECTION			ACTIVATION		
UPDATE PUBLICATION PIPELINE	PRIORITIZE DATA RELEASES	EMPHASIZE QUALITY OVER QUANTITY	ESTABLISH PROGRAM GOVERNANCE	ALLOCATE RESOURCES FOR GROWTH	INCENTIVIZE TEAMS TO PUBLISH DATA	CHANGE CULTURAL PERSPECTIVE	FACILITATE MARKET OPPORTUNITIES	ENCACE LOCAL TALENT	CULTIVATE COMMUNITY PARTNERSHIPS	ENABLE EFFECTIVE USE OF DATA	INCREASE AWARENESS OF OPEN DATA
									ī		
										÷	
P	7	1	P	?	1	P	1	?	ę	1	P
		<u> </u>	Ē		Î		Î		Î	Î	
											¥

# FOUNDATION

Building a strong open data program requires a strong foundation. Current open data publication processes need modernization. To build a strong foundation, the Master Plan will create a transparent development pipeline that will publish more open datasets, and be automated where possible. This theme focuses on quality over quantity by creating higher data standards and new technical requirements that will support a modern and efficient pipeline.



# **Gess** Quality over quantity

# 1a. Update publication pipeline

To enable rapid, comprehensive, and high-value open dataset releases, the City of Toronto needs to focus on establishing an automated data release pipeline that enables internal partners to seamlessly identify, access, and push data to the Open Data catalogue. Decreasing the manual effort required to publish and update open datasets is essential for modernizing and scaling up the City's Open Data program. Real-time access to data in its original, unmodified, and disaggregated form with full transparency on its progress through the publication pipeline is the goal.

#### WHAT WE NEED TO DO

- > Develop an automated publication pipeline that includes privacy and security considerations
- > Make the publication pipeline public for progress tracking
- > Integrate the automated pipeline with the new Open Data Portal
- > Make the publication pipeline self-serve and user-friendly
- > Streamline and minimize executive-level sign-off to publish open datasets
- > Publish real-time open data streams from primary record repositories



#### WHERE WE ARE

- Pipeline is highly analogue, limiting our ability to scale and meet demand
- Datasets sometimes published in a oneoff capacity and often require manual refresh processes
- Point of contact for datasets not always clear or available
- Some datasets originate from secondary sources, reducing transparency

- Publication of the data that powers public web applications on toronto.ca
- Use of the self-serve data publication tools by internal partners
- Effectiveness and efficiency of the open data publication pipeline

## 1b. Prioritize data releases

A process for prioritizing the release of open data with greater consistency, speed, and relevance is needed to meet the growing demands for timely, high value, open data. A prioritization framework will balance demand, publication effort, and value expectations of the data. The prioritization framework, including expected value scores and weighting methodology, will be shared with the public for enhanced visibility into planned open data releases.

#### WHAT WE NEED TO DO

- > Publish the corporate data inventory with regular refreshes
- Identify the publishing status (i.e. In Review, Restricted, or Open) for each item in the corporate data inventory, with estimated release dates
- > Provide a public listing of all open dataset requests
- > Identify where open datasets link to strategic initiatives
- > Target open data releases around key civic issues
- Publish the prioritization framework, including expected value scores and weighting methodology



#### WHERE WE ARE

- Low visibility into release prioritization and publication status reduces transparency
- Corporate data inventory is currently in progress and can be leveraged as a starting point
- Centralized but manual process to request and provide feedback on datasets limits scaling capacity
- Community input on release prioritization not captured via a formal mechanism

- Degree to which open data released enables the City to address key civic issues
- > Gaps in open data around key civic issues
- > Publication of datasets in corporate data inventory
- Alignment between demand for data, data released, and prioritization method

# 1c. Emphasize quality over quantity

Datasets will be scored on quality across various dimensions and relative to other open datasets. Improvements in data quality will be measured and tracked to capture increases in value, and help users appreciate the complexities of data quality. Machine learning will be leveraged to improve scalability, minimize bias, and reduce human error. Data quality algorithms will be published to enhance transparency, reproducibility, and facilitate public input.

#### WHAT WE NEED TO DO

- > Develop a standardized automated quality scoring mechanism
- Create feedback channel for public to engage with data stewards of published open datasets
- Create an internal and external validation group related to the quality scoring methodology and results
- Inform data steward/owners on areas to improve and update data quality at the source
- > Integrate quality score with dataset pages on the Open Data Portal
- Create and develop an open source production environment for machine learning quality evaluations



#### WHERE WE ARE

- Data Governance Framework implementation is underway with a focus on standardization and alignment
- Interdivisional working group formed to advance a common approach to data quality and analytics
- Data quality mechanism enabled via machine learning is currently at early ideation stages
- Dataset quality scoring methodology prototype created for use and evaluation
- Dedicated technology infrastructure for data quality machine learning algorithms in place
- Data quality is a key consideration for publishers yet standardization is very limited
- Methodology for counting open datasets is inconsistent and misrepresents value

- Changes in quality scores for individual datasets and the Open Data catalogue
- Transition of datasets into automated quality scoring process
- > Activity on the feedback channel

# THEME 2 INTEGRATION

To be successful, open data must be embedded within City processes. To do this, open data needs to be a fundamental part of Toronto's government. This theme details the resources required and creates incentives for teams to publish more open data. This theme also strives to change organizational culture to embrace and value open data.



# The default mindset should be we should share this

# 2a. Establish program governance

Governance is critical for maturing the Open Data program. It will lead to increased oversight mechanisms; recognition of open data as a permanent component of the fabric of the city; and instituting policies such as Open by Default.

#### WHAT WE NEED TO DO

- > Create Open Data as an official City program
- > Explore consolidation of open data responsibilities through the office of the City Manager
- Refresh the Open Data Policy and outline updated roles and responsibilities throughout the organization
- Adopt the International Open Data Charter as a policy or as a part of a revised Open Data policy to accelerate endorsement of the "Open by Default" principle
- Form an Open Data Advisory Body with diverse membership comprised of leaders in a number of sectors to champion best practices and enhance accountability
- > Align Open Data program with organizational data governance initiatives



#### WHERE WE ARE

- Open Data is not an official program which compromises permanency, continuity, and consistent access to funds
- Publishing open data is not a requirement and the City's Open Data Policy does not include "Open by Default"
- The Open Data Master Plan Public Advisory Group served as a successful model for a more permanent Public Open Data Advisory Body

- > Adoption of the International Open Data Charter
- Opportunities for public input into the governance of the Open Data program
- Diversity in representation on the Public Open Data Advisory Body
- > Progress in creating Open Data as an official City program
- > Progress implementing the Master Plan

# 2b. Allocate resources for growth

The open data landscape is rapidly evolving and significant investments in human, technology, and other resources are required to meet and adjust to the growing demand. Diversity in skills, shift towards open source technology, and co-development with the community sets the Open Data program for success.

#### WHAT WE NEED TO DO

- Explore creation of a Chief Data Officer role to take ownership of the Open Data program and related areas
- Continue modernizing the current infrastructure to keep pace with evolving industry best practices
- > Fully embrace open source technology and co-development with the public
- > Increase the capacity and diversity in competencies of open data staff
- Establish an accountable open data lead in each Division, Agency, Board, and Commission



#### WHERE WE ARE

- Core Open Data team acts as a single point of contact in open data to Divisions, Agencies, and the public
- Project resources have commenced work on a new Open Data Portal and modernization of the publishing pipeline
- A capital budget for the Master Plan Implementation has been submitted for a multi-year period starting in 2018
- An informal Open Data Community of Practice, made up of individuals across the organization, contributes to advancing the Open Data program

- > Use of open source software in the Open Data program
- > Diversity of competencies within the Open Data team

# 2c. Incentivize teams to publish data

There is an opportunity to enable City teams to better contribute to open data. The Open Data program should provide easy avenues for teams to explore their open data, capitalize on its value through the promotion of their work, enhance collaboration, and track the effort they have put in creating open data for the benefit of Toronto.

#### WHAT WE NEED TO DO

- Improve publisher experience during publication by increasing automation and self-serve capabilities
- Implement simple visualization in the portal to provide partners with basic analysis capabilities of their data
- > Highlight situations in which open data helped address civic issues
- Create a framework for publishers to engage public stakeholders who use their data
- Recognize effort invested by teams and showcase value created across the City
- Introduce internal awards to recognize team successes with open data
- Publish a dashboard that reflects open data contributions by Divisions, Agencies, Boards, and Commissions



#### WHERE WE ARE

- Publishers, at different stages of data maturity, have varying capability to publish open data effectively and efficiently
- Positioning of open data as an approach to decrease FOI requests and repeated public questions
- Shaping value proposition of open data around datadriven analyses and decision-making

- Open data contributions by Divisions, Agencies, Boards, and Commissions
- Effectiveness and adoption of self-serve open data publication components
- Internal use of the Open Data Portal visualization and analytics features
- > Diversity of submissions for the Open Data Awards
- > Inter-divisional usefulness of open datasets
- Creation of data stories published on the Open Data Portal

# 2d. Change cultural perspective

The City needs to remove organizational cultural barriers to open data. This can be fulfilled through recognition of the effort and value groups create within the organization and for the community. Creating safe spaces and strong relationships to address perceived risks in releasing open data needs to be addressed at all levels of the organization.

#### WHAT WE NEED TO DO

- > Recognize collaboration and innovation enabled by open data
- Scale open government course frequency to meet demand and develop a specific open data curriculum
- > Share civic tech community initiatives and stories enabled with open data
- > Enable civic tech community initiatives with corresponding internal programs
- > Establish an internal Open Data Community of Practice
- Provide avenues for internal champions to share stories of open data activation to promote the program



#### WHERE WE ARE

- Successful community engagements have encouraged internal teams to see new value in their data
- Risk aversion, due to concerns like misinterpretation of data, is a barrier to publication
- Sroups have begun to proactively release data
- Internal training courses on open government are available but does not meet high demand
- > Open data is not a priority at all levels of the organization
- Competing philosophies between open data and data for revenue generation

- Requests for publication initialized from the publisher instead of the Open Data team
- Effectiveness and adoption of training courses in the open data curriculum
- Depth of membership and engagement in the Open Data Community of Practice

# THEME 3 CONNECTION

Toronto's thriving and diverse communities are central to the open data movement. The ecosystem of open data users is growing and so are the opportunities to create value. This theme recognizes that partnership development is mutually beneficial and necessary for achieving the vision and goals of a mature open data program.



# Allow Torontonians to help City staff build for the public

# 3a. Facilitate market opportunities

Economic development and vitality in the digital era requires access to current, quality, relevant and authoritative data. The City has data and information that could be released as open data to enable entrepreneurs and private organizations to develop market viable solutions. New and relevant services created by third party organizations not only benefit the community, but also stimulate the local economy.

#### WHAT WE NEED TO DO

- > Capture knowledge on how open data is used throughout multiple sectors
- > Create mechanism for users to sign up for notifications when datasets are updated
- > Identify potential of unreleased data to unlock market opportunities
- > Increase engagement with local business to understand how open data fits in business models
- > Ensure open data is available in the formats and channels best suited for business use
- > Educate entrepreneurs and startups on open data and key civic issues
- Work with City subject matter experts to understand how open data is contributing to the business sector
- > Implement a targeted market profile of open data business consumers



#### WHERE WE ARE

- Datasets published in various formats, including development-friendly APIs
- > Data provided is beneficial to private businesses
- Minimal visibility into how open data is being used by local businesses, entrepreneurs, and startups
- Limited understanding around what market opportunities can be unlocked by opening specific datasets
- Regular contact with Open Data Exchange (ODX) provides insight on how open data is utilized in the innovation and business sector

- Adoption of notification system
- Solutions powered by the City's open data
- Monitor changes to market profile of open data business consumers

# **3b. Engage local talent**

Toronto, the Greater Toronto Hamilton Area (GTHA), and the Toronto-Waterloo Corridor, enjoy a high concentration of unparalleled talent, academic institutions, and entrepreneurial culture that could be leveraged to identify opportunities and address challenges facing the city. Open data can act as a conduit between these groups, serving to foster an ongoing relationship designed to improved civic, social, environmental and economic vitality for Torontonians.

#### WHAT WE NEED TO DO

- > Become actively involved in community-driven groups and organizations related to open data
- Position civic issues facing the City as challenges community groups could help address with open data
- > Release datasets in the Open Data Portal that are tied to key civic issues
- Host events for application developers, data scientists, researchers, and others to collaborate on solving civic issues important to the City
- > Contribute to an increased and wider variety of hackathons and data-driven events
- > Capitalize on emerging sectors in Toronto's tech community, such as AI and Machine Learning



#### WHERE WE ARE

- Regular participation in community-driven activities around open data
- Ad hoc hosting for civic tech community and related events
- Occasional open data team participation in local hackathons
- Aware of local developers and researchers interested in open data due to ongoing community outreach

- > Engagement level in events related to open data
- Progression of community-centric solutions for addressing civic issues

# 3c. Cultivate community partnerships

In order to tackle the civic and social issues facing Toronto, the City cannot work in isolation. Greater community involvement requires a focus on establishing, fostering, and growing a network of mutually-beneficial community partnerships that will produce benefits for residents of Toronto.

#### WHAT WE NEED TO DO

- > Provide access to City subject matter experts (SME) on a range of key City civic issues
- > Transition community connections into formal partnerships with specific goals
- > Provide greater opportunities for community organizations and startups to partner with the City
- Create a mechanism to share and pool resources that result in actions related to solving civic issues
- > Partner with various City divisions and programs that will enable effective community partnerships



#### WHERE WE ARE

- Initiated and fostered connections with various community organizations and leaders related to open data
- Engaged with Provincial, Federal and other municipal open data teams to share best practices, lessons learned, and coordinate some strategic initiatives

- Use of open data for helping solve civic issues through official partnerships
- Level of collaboration between the City and community partners focused on civic issues

# THEME 4 ACTIVATION

The value of open data will be best realized by allowing a greater variety of users, both technical and non-technical, to have equitable access to open data. This theme outlines how investing in opportunities for enhanced digital literacy, and increasing understanding of how government operates, will provide the neccesary context to acitvate open data.



# Tell us what you want versus tell us what you want you've got

# 4a. Enable effective use of data

To realize the value of open data, Toronto needs to shift focus from solely a catalogue of open datasets targeting the developer community to a platform that makes open data accessible to a wider audience. Novel forms of engagement and dataset exploration will enable greater activation of open data, grow value creation, and meet the responsibility for inclusivity. A principal focus will be helping Torontonians develop the knowledge and tools needed to make use of data.

#### WHAT WE NEED TO DO

- > Increase accessibility of open data to a wider audience with a range of technical abilities
- > Provide context around open datasets to enhance understandability for a wider audience
- Develop and embed inventive open data development models like "Open Data In Reverse" for more effective data releases
- Position open data within emerging City frameworks like the "Insight Activation Cycle"
- > Provide on-demand data format conversion
- > Promote community sourced and 3rd party datasets
- > Provide guidelines for the creation of community and 3rd party sourced datasets
- > Implement visualizations of open datasets in the portal
- Create open data education/instructional videos

# 🗗 🔅 📍 💿 🏛 🍟

#### WHERE WE ARE

- Existing Open Data Portal contains a dataset catalogue and showcases select applications and visualizations created from it
- Open datasets available in multiple formats to meet the needs of a wide variety of users, but format availability is inconsistent across datasets
- Developing a new Open Data Portal based on usercentred design to capture user feedback
- Created prototypes for data visualizations and stories to give data context and test ways to make it more widely accessible
- New Open Data Portal will be made available for internal and external evaluation
- Situated open data within frameworks related to data governance and activation so the City can harness the value of open data

- Increase accessibility of open data to wider audience with a range of technical abilities
- Publication of community and 3rd party open datasets
- Usage trends of portal features and data formats
- > Engagement levels of target communities
- > Utilization of open data education tools

# 4b. Increase awareness of open data

Greater awareness of the benefits and enabling factors of open data is needed, internally and externally. Increasing digital literacy and providing opportunities to understand how government works will be key drivers to unlock the potential of open data.

#### WHAT WE NEED TO DO

- > Identify how open data is used as part of data-driven decision making
- > Share open data success stories across the City and community
- Establish City Awards to recognize external contributors to civic-focused initiatives powered by open data
- > Develop approach to showcase value of open data
- > Deliver Open Data 101 courses
- > Build understanding around the program related to open data releases



#### WHERE WE ARE

- Open data awareness has increased in the City, but a stronger grasp of its value is still needed
- Ongoing awareness efforts center around a highlyengaged community consisting primarily of developers
- Starting to release open data in tandem with City initiatives and programs

- > Dataset consumption
- > Usage trends of datasets internally and externally
- > Data stories published on the portal
- > Participation in Open Data 101 courses

An implementation roadmap is an essential element of the Master Plan and highlights areas of activity and their sequencing. It provides a high-level direction for activating the plan and ensuring tangible outcomes for the City's Open Data Program. The roadmap should be flexible and adjust to needs and evolving opportunities.

# How we will get there





### How to get engaged

Explore materials created
through the development
of the Open Data Master
Plan

**PUBLIC FOLDER** goo.gl/gj1rTg Leverage and contribute to the City of Toronto's public code repository

GITHUB REPO github.com/CityofToronto Share how you use City of Toronto open data

**@OPEN\_TO** opendata@toronto.ca

Meet and chat with us at community events

Stay up to date on what we are doing

OPEN DATA BOOK CLUB opendatabook.club **CIVIC TECH TORONTO** civictech.ca

**OPEN DATA PORTAL** toronto.ca/open Open Data Master Plan \_\_\_\_\_