035

North America, Australia, and New Zealand

David Eaves, Ben McGuire, and Audrey Carson

Key points

- The shuttering in 2017 of open.whitehouse.gov illustrates the risks that face open data movements in Australia, Canada, New Zealand, and the United States (US) as high-level leadership on open data and access to certain datasets are vulnerable to changes of administration. However, open data practice at local levels has proven reasonably resilient so far.
- The last decade of open data work has encouraged governments to perceive their data as an asset that needs to be managed and to improve internal capacity on data use and analytics. In particular, open data has played an important role supporting cross-agency collaboration.
- Direct citizen demand for data from national open data portals remains low and highly concentrated. This can leave initiatives struggling to define success when they operate on the simple logic that open data will boost citizen engagement.
- More work is needed to both define what success looks like for open data initiatives and to monitor progress toward it with the right kinds of tools.

Introduction

One of the Trump administration's early decisions after taking office in 2017 was to shutter open.whitehouse.gov, a portal set up by the Obama administration to house public information on visitor logs, financial disclosure reports, and other personnel data. The White House Communications Director at the time, Michael Dubke, suggested that the move was in response to "the grave national security risks and privacy concerns of the hundreds of thousands of visitors annually".[1] To open data advocates, the disappearance of this data was an early sign that "American leadership on open government will not come from this Presidency".[2]

And while the closure of open.whitehouse.gov could serve as a reminder of how fragile gains in open government can be, the reality is that federal rules and legislation, like the DATA Act, continue to build and preserve open data sources across government, and the administration has publicly committed to improving the way it uses and supplies data to the nation.[3] The events of last two years also suggest that, even without executive leadership, officials at the state and municipal level have generally been able to continue to develop open data activities within their own areas of work. The past decade has been an extremely exciting time for open government and transparency in the United States (US), Canada, Australia, and New Zealand. Across these nations, there are powerful stories about open data initiatives that demonstrate a greater government commitment to, and the potential benefits of, making information more transparent.

But despite the robustness of open data initiatives during the transition to new governments with different political stripes across North America, Australia, and New Zealand, the news is not all positive. Participation remains extremely unequal across national and subnational jurisdictions. While the number of datasets on federal open data portals initially proliferated as agencies scrambled to comply with national standards, the usability and relevance of this data is often still in question. Early success stories related to geospatial data, metropolitan transportation, and legislation-tracking applications have not translated to success across all geographies or sectors. Instead, pockets of urban innovation continue to dominate the demand for, and supply of, open data, while the participation of local governments remains vulnerable to political crosswinds.

Even with these gaps and challenges in mind, the public servants, open data advocates, and private sector partners surveyed for this chapter continue to find reasons to be cautiously optimistic about the present and future of open data. While the use of open data by the general public lags behind the most optimistic predictions, usage by government staff and intergovernmental cooperation is increasingly becoming a driver of open data's value. Whereas early efforts focused on the volume of open data produced as a measure of success, governments are now becoming more sensitive to tracking usage and demand, and taking a more user-centred approach to future reforms.

Also worth noting is that the incorporation of open data as an expected practice at all levels of government has helped to build the capacity for proactive data management and analytics. Indeed, it could be argued that the open data movement's biggest accomplishment may not be the raft of new applications created by third parties, but rather the fact that it has prompted governments to perceive their data as an asset that needs to be well managed to be effectively leveraged. Jurisdictions that adopt open data aggressively tend to also manage and use data more effectively, understanding its implications on issues such as privacy, as well as how it can be used to improve operations. Open data adoption may have given some governments a jumpstart on digitising and modernising their governance.

This chapter lays out trends in open data across the region over the last decade, focusing closely on signals of progress, persistent gaps, and the potential long-term analytics value of open data investments. While powerful and positive anecdotes reflect the strength and creativity of the open data movement, due to challenges in defining shared goals, maintaining sustained investment, and matching the supply of datasets to citizen demands, stagnation and drift might better characterise the current general state of open data in this region.

Signs of progress

In 2013, the Government of Canada officially re-launched its open data portal, and, in the US, President Barack Obama signed an executive order making open and machine-readable the new default for government information. These landmark federal actions came just five years after the City of Washington in the District of Columbia launched its open data portal (2007), leading to an explosion of open data portals that included fast followers like Vancouver (2009), San Francisco (2009), New Zealand (2009), New York City (2012), and Chicago (2012), among others.[4] In the years since these first major open data breakthroughs on the North American national stage, public servants, citizen hackers, and software companies have all played a role in driving spotty but real progress in the world of open data.

To fully assess the recent progress of the open data movement across North America, Australia, and New Zealand, interviews were conducted with open data leaders across the public and private sectors. Among the participants were a state Chief Data Officer from the US, an Australian provincial Information Commissioner, government practitioners at the federal and local levels tasked with implementing open data directives and scaling open data portals, and the directors of well-known civil society organisations and academic research groups specialising in open data.

Taken in aggregate, their accounts reveal a qualified optimism about the present status and future trajectory of the open data movement. While most believe that progress has been harder and slower to come by than expected, nearly all were also able to point to major milestones that have been reached, as well as other compelling instances of progress. They universally agreed that the adoption of open data standards has gradually broadened, and that open data programmes are beginning to yield positive public benefits from greater government transparency to improved bureaucratic efficiency. Due to the efforts and investments of these open data advocates, as well as other committed individuals and organisations, it is evident that slow but steady steps have been taken to increase the ubiquity and usefulness of open data in a number of jurisdictions in these regions.

Open data leaders emphasise different goals when describing what drives the open data movement, whether it be ensuring government accountability, spurring private sector innovation, or improving bureaucratic decision-making by breaking down barriers to the access and reuse of public information. It is heartening to see that in specific instances, open data programmes have made headway against all of these goals. Highlighted below are a few anecdotal examples of the positive impacts that policies and products, fueled by open data, have had within different sectors or open data communities:

Transportation – Some of the most visible applications to date have been in the field of transportation. In New Zealand, ANZ Bank has taken advantage of open data on the traffic flow of trucks and heavy vehicles in order to forecast GDP, which is published monthly under the "Truckometer" name.[5] Another example of open data at work in the realm of transportation comes from the Australian province of New South Wales, where developers have pulled data from the province's open data portal on the location of petrol stations and station-specific petrol pricing to provide consumers with a transparent look at where they can be cost effective when

filling up their tank.[6] In fact, practitioners have indicated that this application has helped prevent petrol price gouging across the province and encouraged more consistent pricing across stations, producing a net benefit for consumers. The increase in accessible data has drawn new actors into the space, such as Google with Google Maps, as well as new companies, like Remix, which provides an advanced planning tool that transit agencies can use to test transit variables and schedule complex transit services.[7]

Health – The US State of Connecticut is at the forefront of utilising publicly accessible data to monitor threats to public health and make data-driven adjustments to policy and programmes to better serve the needs of its residents. In particular, the state has used open data as a tool to help fight the opioid epidemic that has seized the US in recent years. Using data about the location of pharmacies that prescribe opioids, dispense Naloxone (a medication used to reverse an opioid overdose), and provide prescription drug drop boxes, as well as opioid overdoses and related deaths, the state has been able to trace the movement of the opioid epidemic from one neighbourhood to the next. These findings, in turn, have been used to prioritise treatment for opioid addiction over treatment for other conditions at treatment centres in certain heavily affected neighbourhoods and to more effectively stock pharmacies located in the epicentre of the epidemic with Naloxone in order to avoid life-threatening drug shortages.[8] Some US municipalities have also been able to utilise a similar cross-departmental approach to solve less severe, yet still entrenched, public health problems, such as the disproportionately high (and expensive) utilisation rates of emergency services by a relatively small portion of the population. The town of Cary, North Carolina has attempted to solve this problem by aggregating data across agencies from health, law enforcement, and other social services to identify root causes and provide better preventative, non-emergency care to these high-need individuals, thereby reducing the overall cost of the health system. Indeed, by investing a relatively small sum (around USD 200 000) to publish this data and conduct the relevant analyses, Cary was able to make adjustments that local practitioners estimate now saves the city more than USD 8 million annually in health costs.

Environment – The boreal forest is central to Canada's natural environment, history, and culture, but it is also an important economic resource that is regularly harvested, contributing significantly to the country's overall GDP. The 2010 Canadian Boreal Forest Agreement was a landmark conservation initiative that brought together industry, environmental groups, and government to broker a consensus on forest management techniques in the region. Central to the agreement was an accord to prohibit logging in certain vulnerable forest areas and adopt data-driven sustainable harvesting practices across the region.[9] In return for these protections, Canadian environmental groups committed to no longer boycott forestry goods produced by the participating lumber companies, which increased the value of their stock by stabilising consumer demand. Much of the data on forest health used throughout the negotiations was open data provided by Environment Canada, such as woodland caribou range data. By serving as a single source of truth throughout the negotiations, Environment Canada provided a common foundation for both sides when debating the best way to integrate economic and environmental values. Without this open data, it is likely that the accuracy and sophistication of the debate would not have been

possible and that an agreement would not have been reached. In this case, open data was a foundational ingredient for two competing interest groups to engage in productive, evidenced-based policy-making.

Law enforcement – Across these nations, the utilisation of open data by police departments stands out as a particularly high-potential and effective application of open data for public benefit. In the US, the Police Data Initiative, which started in the White House of President Barack Obama, but continues under the stewardship of the Police Foundation, provides a consolidated and interactive listing of open and soon-to-be-opened datasets identified by more than 130 local law enforcement agencies as important to their communities.[10] In this context, open data is used to encourage joint problem-solving, innovation, enhanced understanding, and accountability between communities and the law enforcement agencies that serve them. Most recently, the organisation has called on local agencies to publish data about hate crimes and has received a promising response from heartland American cities in Nebraska, Oklahoma, and Indiana. Other local police departments, such as the Toronto Police Service, have launched their own open data sites to share data directly with the public.[11] The adoption of a publish-by-default approach, as is the case with New Zealand's police agencies,[12] is a truly important development in the evolution of the open data movement, representing an impressive maturation in the way local agencies engage with open data. Notably, regular publication of New Zealand Police's crime statistics has led to detailed investigative journalism in both print and online media.

Elections and government accountability – Election data has proven to be one of the most popular types of data published on open data platforms. For example, Elections BC, a non-partisan office of the British Columbia legislature that is responsible for conducting all public elections in the province now releases all of its election results as open data.[13] Small municipalities, like the Township of Langley, Canada, have also found that election-related data is among the most heavily accessed data via their open data portals. Independent initiatives, like OpenSecrets.org produced by the Center for Responsive Politics, have also meaningfully expanded citizens' access to in-demand campaign data. For other open data applications, the aim is not so much to document the outcomes of elections as it is to empower citizens to hold government accountable at the ballot box. Some civic hackers in New York City, for example, have used public utility data revealing insufficient heating in the city's public housing projects to advocate for new leadership at the local and state level, ensuring proper funding of the New York City Housing Authority.[14] Other open data applications in the US, like GovTrack[15] and Councilmatic,[16] have helped to increase civic engagement by making it easier for citizens to stay up-to-date on their elected representatives' voting history, as well as pending legislation and special events.

While the developments detailed above from across this region are diverse, taken together, they speak to a number of common trends underlying the state of today's open data movement. First, increased collaboration on open data across government units, such as when a public health department shares anonymised data with a police department when both are tasked with treating the same problem, has allowed for unprecedented analytics and insights into the complex causes of various public policy outcomes. Second, publishing more and better data has enabled the

development of applications that have netted measurable benefits for citizens and residents. Citizens and activists who have the ability to easily access vast quantities of government data are empowered to more accurately assess and act on information in both their private and public lives.

Above and beyond these trends, however, two developments stand out that seem likely to bend the arc of the open data movement by accelerating its adoption and impact: the move toward internal usage and the move toward quality over quantity.

Internal usage

Among the first wave of state and local open data portals, the primary objective was to publish anything at all, regardless of whether it would actually be accessed or not. In this environment, releasing data was driven by a need to comply with the requirements of overarching data directives, not by the specific uses and desires of downloaders. Over time, however, government departments and agencies tasked with publishing open government data began to refine their criteria for success, adopting goals related to the external usage of data by citizens. In this second wave, success was more often assessed via key performance indicators that measured the number of datasets downloaded. With limited means to assess how data made available was being used, government entities focused on sheer citizen demand for open data to justify their programme costs. Underlying this operating model was a primary objective to make government data available outside of government for use by citizens.

In recent years, however, there has been a shift within the departments and agencies that publish open data. This shift has been driven by a need to rationalise the costs of maintaining an open data programme and to convince internal stakeholders and bureaucrats of the practical value that open data programmes can provide for the day-to-day delivery of government. Today, open data portals are increasingly designed to maximise the usage, sharing, and analysis of open government data by government employees themselves, with the end goal of supporting evidence-based policy-making and improved performance management.

Under this new operating model, the value of the open data programmes as a whole can be measured partially by the incidence of government employees actually utilising the data to make data-informed decisions, both when developing policies and when adjusting existing policies to reduce costs and improve outcomes. While a mandate to "defend the spend" may not sound as lofty as proclamations to advance open data as a transparency tool, the shift toward internal usage is a truly positive development. Justifying the value of open data portals internally helps to secure their long-term sustainability, ensuring that both government employees and the public are empowered to extract meaningful value from open data in the years ahead.

Quality over quantity

The open data practitioners interviewed for this chapter frequently cited a struggle between the desire to publish a large quantity of data and the desire to publish high-quality data. In the early years of the open data movement, efforts to rapidly scale up open data programmes, either in response to real user demand or in order to match or exceed benchmarks for success, had the unintentional effect at times of encouraging the publishing of non-standard or non-machine-

readable data. While the quantity of datasets published on open data portals increased rapidly during this time, actual utilisation of this data often stagnated, hobbled by the data's overall low quality and usability.

Today, governments like Canada's lead the pack in a larger effort to shift the emphasis from sheer quantity of data to the quality, consistency, discoverability, and coherence of data. To this end, governments have begun publishing data inventories and indexes to assist with data discovery, organising datasets to allow more coherent analysis, and developing additional user tools, such as data maps and visualisations, to help laypersons better understand the data hosted on their platforms.

Canada's Information Management and Open Government team has even gone so far as to suspend the practice of including the number of datasets available in reports on the health of the country's open data programme because this metric has little bearing on the platform's utilisation and overall success. Indeed, in recent years, the number of datasets available on Canada's open data portal has dropped from around 300 000 at its peak to approximately 80 000 today. This decline is a testament to the government's commitment to painstakingly auditing the datasets on the platform and replacing non-contextual, non-machine-readable data with more consolidated, accurate, and digestible data resources that will actually be used.

Emerging challenges

Alongside the trends above that demonstrate a maturing of the open data agenda across North America, Australia, and New Zealand, initiatives are also facing a number of challenges.

Slowing or stagnating data publication

Despite the many signs of progress and anecdotal success identified in the preceding section, a dominant narrative on open data at the end of its first decade is one of stalled progress, in which both advocates and practitioners are forced to reassess early expectations in light of unexpected technical, political, and cultural limitations. Early ambitions for rapid adoption fuelled by innovative open data applications in venues like New York City subways or gas stations in New South Wales have given way to a recognition that further gains will not come simply by increasing the supply of data or enacting legislation. Despite optimism about the future of open data, one practitioner in the US related a common thesis that "open data has become a bit stagnant here as we continue to focus on broader adoption and more open data instead of on what data may be the most valuable in a national context".[17] Rather than increasing transparency and the uses of open data year over year, these countries are still struggling to maintain and drive further adoption and compliance.

One insight into this trend can be found by looking at comparative international data from the Open Data Barometer.[18] Examining the results of country-level surveys,[19] it is possible to explore how 16 data categories are rated on a scale from 1 to 100 based on standardised factors, such as openness, including open licensing, machine readability, cost, and timely updating. Since 2013, several of the data categories have obtained consistent high ratings (e.g. mapping, trade, elections), while others rise (e.g. company) or fall (e.g. spending) due to issues related to outdated

or inconsistent published data. These variations aside, the overall picture is one of relative stability (see Figure 1).

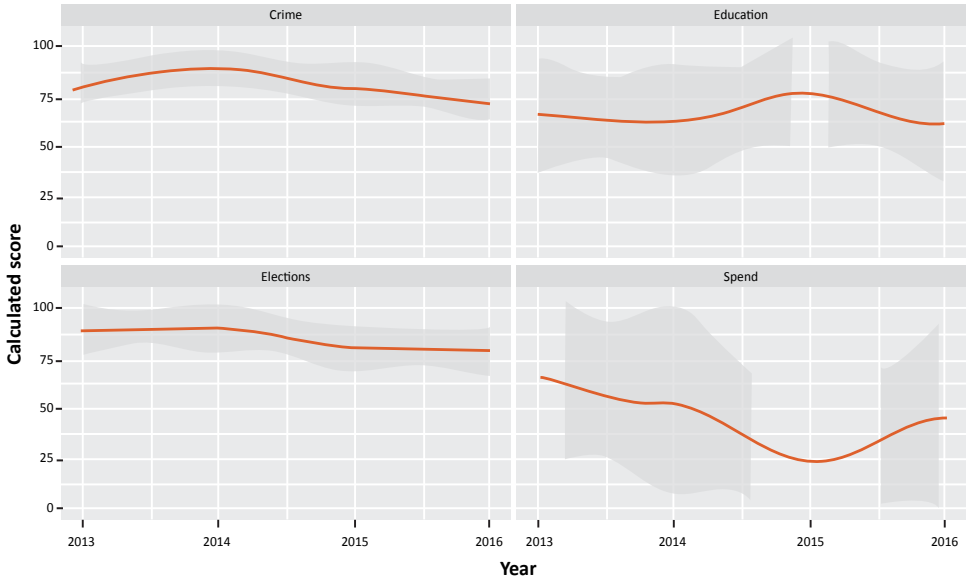

Figure 1: Average regional sector openness scores (2013–2016)
Source: Open Data Barometer 2017, consolidated data: https://opendatabarometer.org/4thedition/data/

Stagnation does not only plague the general state of data openness and access. A closer look at national open data portals in North America, Australia, and New Zealand indicates that a slowdown has also occurred in terms of the data being uploaded to flagship sites. While advocates rightly celebrate the increasing participation of various agencies within governments in the release of open data, the broader story is one of early adopters and laggards. By and large, the picture today is very much the same as it was five years ago.

In the US, the national open data portal Data.gov collects hundreds of thousands of datasets into a single, searchable database. With over 365 000 total datasets uploaded to date, the federal government has been by far the greatest contributor to the Data.gov repository, accounting for 85% of the total datasets. State governments account for only 5% of the total. Within organisations, the distribution of datasets uploaded by individual offices can be extreme. On the low end, 50% of all offices have uploaded five datasets or fewer. At the high end, three federal agencies (Commerce, Interior, NASA) account for over 60% of all datasets on Data.gov.

The temporal pattern of uploads to Data.gov over the last decade also tells a story of inconsistency and occasional spikes of activity linked perhaps to major policy announcements, which are followed by limited ongoing participation. While the metadata on site uploads over time is not available for all countries within the group discussed here, practitioners interviewed

for this chapter suggest that a similar pattern of stagnation in agency participation on open data portals is a significant problem within and across their home countries.

It is not clear, however, whether this represents the demise of open data initiatives or simply the demise of the open data portal as the organising device within such initiatives. As agencies shift to focus on data for targeted groups of users and to internalise open data practices, they may be making more use of their own websites to distribute datasets, providing specialist data platforms instead of using national portals. For example, in New Zealand, geospatial data is provided through the LINZ Data Service,[20] which users may first discover via the national data portal (data.govt.nz), but once found, it becomes the more direct destination for geospatial data.

It is clear that national open data initiatives in the region have, in general, been experiencing a sense of disillusionment that characterises many new technologies that move beyond the peak of a hype cycle, a challenge that other nations across the world may soon have to grapple with. According to the open data advocates interviewed for this chapter, this slowdown in high-level participation is rooted in four distinct challenges:

- The struggle to agree on definitions for goals and success.
- Uneven adoption within and across government at all levels.
- A reliance on champions who do not maintain their roles over the long term.
- A focus on performance metrics related to supply rather than demand and use.

How these challenges are addressed will shape the development of open data in the next decade.

The struggle to define success

The public presentation of open data and the stated aims of public programmes frequently trumpet transparency and the potential for civic engagement. However, many actors also believe that the potential economic benefits from sharing data as a "natural resource" will increase the efficiency of government services or facilitate new opportunities for economic growth, essentially meaning that open data programmes would pay for themselves. The potential tension between the "good government" and "economic benefit" perspectives can sometimes lead to confusion when trying to prioritise which datasets to share and, in the medium term, can make it difficult to agree on how to define programme success.

In the early stages of the open data movement, the US and Canada stressed transparency and civic engagement as the primary motivations for making information more available. At the opening of Data.gov, the Sunlight Foundation's Executive Director referred to the launch as "a dramatic breakthrough in the role of government".[21] In New Zealand, government's framing of its open data initiative was broader, linking open data with open government agendas, but also maintaining a degree of distinction between the two. Regardless, many arguments for open data focused on its broad potential as a resource for change without setting out explicit measurable policy goals. This constructive ambiguity in defining the goal of open data initiatives and the

subsequent way in which different groups have engaged with open data, has, over time, made the identification of success or failure a significant challenge.

According to practitioners interviewed for this chapter, many early adopters primarily understood open data as a "transparency tool" that would provide citizens with "more access to the inner workings of government" while driving new "civic participation and accountability in the public sector".[22] Early experimental and creative apps that took advantage of new datasets promised an explosion of citizen data use, but in the decade that has followed, such a transformation has yet to materialise. Instead, as government budgets have held steady or shrunk, there has been increasing pressure to defend open data as an investment and to define a return value. Even ardent advocates admit that transparency is "hard to quantify, to implement, and defend from a budget perspective".[23]

As US and Canadian narratives shifted away from transparency justifications and moved toward discussions of open data as an asset and a tool for making government more efficient, practitioners tried to drive value by uploading lots of data in the hope that beneficial uses would organically emerge. As one practitioner described this period, "We put up what we thought people wanted to see."[24] For advocates and practitioners, the shift to prioritising economic value raised the immediate expectation of revenue or savings from open data, but these also were not necessarily forthcoming or easily quantified in the short term. Ultimately, it is difficult to "make the change and articulate the overall benefit in the long run of early government investments, which will reap tax revenues or efficiencies from uncertain new ventures".[25]

One dangerous consequence of this lack of agreement over goals has been a reversion to competition and comparison to other governments, rather than increasing the focus on measuring the value of local investments by the good they do for consumers and citizens. Programmes like the Open Cities Index[26] or the Open Data Barometer[27] were never intended to be the final definition of success, yet some municipal open data practitioners are forced to define their success via these tools even when they "don't really understand the criteria used to judge",[28] because they impress elected officials.

Uneven adoption within and across governments

While inter-agency and inter-governmental collaborations are exciting developments in some areas as noted earlier in this chapter, a sizable challenge to broader engagement with open data and sustained citizen involvement is the difficulty that many governments still have working together. This problem is particularly acute at the subnational level (state and provincial), where there may not be clear consumer demand for data due to a lack of public awareness or lack of demonstrated use cases. In contrast, for large cities, the density of the population and volume of data produced, especially for services like transportation, has allowed for sustained buy-in, but, beyond questions related to scale and opportunity, there remain significant data quality, governance, and definition challenges, which inhibit progress.

Many practitioners interviewed for this chapter suggested that the open data movement has struggled because there has not been enough attention paid to interjurisdictional collaboration and data standardisation. As a result, open data plans at the national level tend to dominate the agenda, while the coordination between levels of government, which practitioners believe could

produce the greatest value, is often forgotten. This challenge manifests not only at the federal level between different agencies, but is also a significant problem where regional and municipal governments have begun to set up open data federations. While these cooperative groups represent one of the largest current opportunities in the open data movement, they face huge hurdles in coordinating datasets across levels of government and with private sector partners. One of the complex barriers to cooperation is data ownership, particularly when private data is involved, or when both public and private sector organisations are involved in the same project. However, where cross-sector data co-operatives are able to form, such as the Sault Ste. Marie Consortium in Ontario, Canada, which has 22 organisational members of different sizes,[29] they are able to leverage skills from across the partnership to share skills and generate analysis based on shared data.

At the state and local level, practitioners report that a simple shortage of skills prevents them from reaching the sort of scale and value realisation that is frequently identified in large cities. One practitioner related a common lament that there is still a "long way to go with smaller local governments, because they don't have the political support or capacity to move on their own".[30] A huge part of the challenge in this area is simply based on volume and demand. There are few organisations or individuals seeking public data from states and small cities relative to very large metropolitan areas because the minimal volume of data produced there does not produce as many obvious opportunities. Until these regions can cooperate more effectively or find ways to drive increased demand, it is likely that subnational open data movements will continue to struggle outside of the largest cities.

Reliance on local champions to build and sustain initiatives

The ultimate goal of open data advocates and practitioners is that data will become "open by default"; however, no governments have reached this goal as a steady state, and many still closely guard access to most of their datasets. In some cases, this may be because "authoritative publishers, particularly in government, are paralysed by a fear of publishing poor quality or incorrect data"[31] for fear of damaging any existence of trust with constituents. Consequently, practitioners admit that, in many cases, "open data programmes need to have very strong-willed champions or demonstrated return, or else they'll wither off".[32] When political turnover or other factors result in the removal of these key personalities, open data programmes can falter or even disappear.

Virtually all of the practitioners and advocates interviewed for this chapter suggested that open data movements in the context of their own experiences could not have succeeded without a handful of strong local champions. However, few could articulate an effective way to hand off the administration and maintenance of programmes to the next generation of potential leaders, and many admitted that open data remains very susceptible to budget cuts that result from political turnover. The challenge is distinct within and across the countries discussed in this chapter. In New Zealand, for example, public service chief executives are appointed on five-year contracts independent of political change, unlike their counterparts elsewhere in this group of nations. Personality-driven projects can result in open data initiatives that are siloed in one part of an organisation, and, if open data is seen as an IT project rather than a strategic asset for example, long-term sustainability is a challenge.

Citizen demand for data remains limited and highly concentrated

The most consistent gap cited by those interviewed for this chapter related to the measurement of outcomes and, in particular, how to best track the relationship between open data producers and potential users. In the earliest years of the open data movement, the release of large amounts of open data represented a kind of progress, but practitioners now suggest that creating more datasets "doesn't allow value to be realised" because producers have largely "failed to create an authentic feedback loop between users of data"[33] and government.

One way to explore this problem is by looking at usage metrics from national open data portals to see not only how the demand for data has changed over time, but also what total demand looks like relative to a nation's population and how concentrated demand is for specific datasets. When per capita uptake is low, this could suggest that open data in general has not been adopted by the public for mainstream use. When use is highly concentrated among a handful of consumers, this could suggest that rather than serving a diverse set of interests, open data tools are essentially catering to a specific power user base.

In the US in 2016, there were about .009 recorded views of datasets on data.gov per resident with internet access. In Australia, the equivalent figure for 2017 was .14, and in Canada, it was .04 (New Zealand does not make the metadata required to calculate this ratio available). This data suggests that larger countries may suffer when judged by the volume of completed downloads (and perhaps also hints that the size of an open data community is not directly proportional to nation size); however, it could also suggest that approaches to advertising open data portals to the public have had different levels of success within different countries.

We can also use portal metrics to explore trends in the uptake of open data. A first glance at monthly visits to Data.gov suggests a noisy but discernible upward trend in site visits (see Figure 2 below).

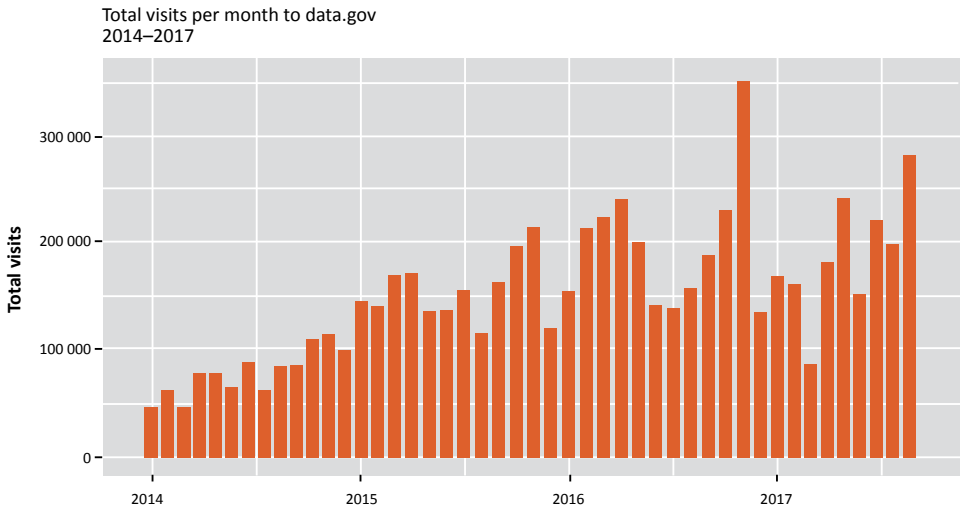

Figure 2: Monthly visits to data.gov (2014–2017)

A closer look at the monthly visits indicates that there is significant variation within and across years, but that most datasets remain consistently below 2 500 monthly views. In fact, most agencies have had an effectively static rate of page views since 2014. Instead of showing a general upward trend, the data suggests that overall increased traffic stems from a few outlier datasets that have received much higher traffic. Notably, the ten largest agencies by dataset count (about 4% of agencies releasing data) account for 75% of total downloadable files and 40% of total views.

In Australia, a similar upward trend emerges over time (see Figure 3), providing an encouraging sign that more citizens are taking advantage of the portal. However, as with the US data, examining publisher popularity on data.gov.au shows that visits are dominated by a small handful of agencies. Datasets published by the Department of Human Services alone account for 47% of all visits to the site since 2013, and the ten most popular publishers account for 84% of total visits. By far, the most popular resources on data.gov.au are a small sample of location search resources, with the locations of Medicare offices accounting for 27% of total historical visits and the Location of Centrelink Offices accounting for another 14%. Essentially, two out of 70 000 datasets account for 41% of all visits. The distribution of downloads rather than visits is slightly less skewed toward the most popular resources, with the ten most popular datasets accounting for about 54% of visits, but only 21% of downloads.

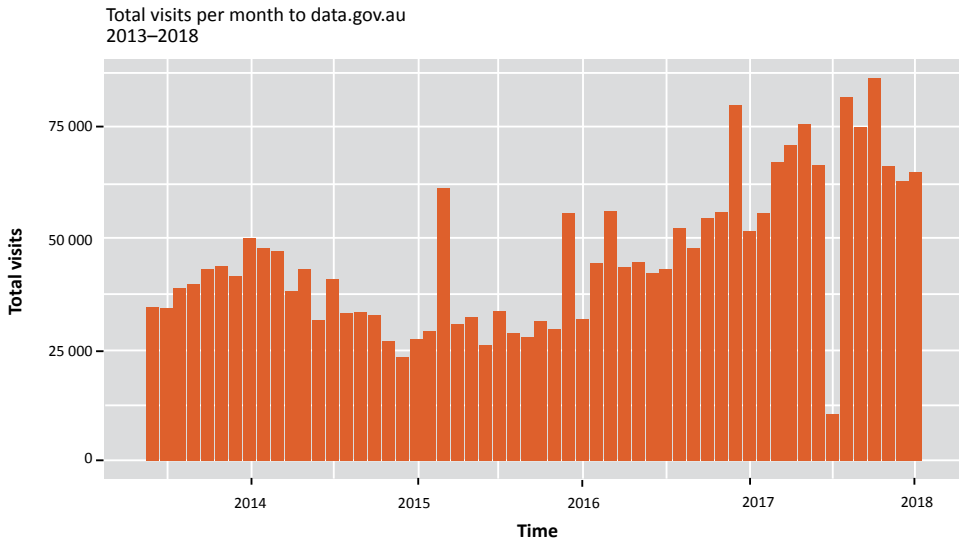

Figure 3: Visits to data.gov.au

Other countries in this group do not provide this kind of data, but additional analysis and anecdotal evidence suggests that demand is similarly highly concentrated on a small percentage of available data, indicating that much of the information that has been made public does not serve any specific consumer need. Should administrators interpret the gap between what users are accessing and the vast resources available as a problem that needs to be solved? Some practitioners suggest that this skewed distribution was an inevitable result of experimentation

and reflects the fact that many users are simply not aware of what is available. Others argue that there are legitimate "difficulties in identifying who currently uses open data, who might use it if certain changes were made, and how all users think about the reliability and usefulness of government datasets relative to other tools".[34] Even the best anecdotes about users and their open data success stories rarely translate to strategies for how to promote the use of other data. Establishing lines of communication with users that can reveal their preferences can be difficult and time-consuming when both data and user needs change over time.

Hidden benefits: Analytics and performance management

One potentially underappreciated aspect of the open data story is that beyond the supply and demand for transparent government information, open data may be laying a broader foundation in the region for more robust analytics and performance management in government.

Generating and making sense of open data forces staff to develop skills to manage, analyse, visualise, and interpret complex datasets. Furthermore, investments in open data elevate data managers and analysts into strategic discussions and help enable nominally operational teams to take on strategic roles. Open datasets have formed the foundation of some of the most promising public sector innovations in predictive analytics,[35] and, in many cases, large cities in the region that are leaders in data-driven governance have built up their expertise based upon earlier work with open data.[36]

From a management perspective, increasing organisational capacity for discussing and interpreting data helps develop a culture that can be more intentional about establishing shared definitions, defining success through metrics, and creating incentives that rely on common datasets. None of this obviates the critical role of progressive and aggressive leadership, but practitioners argue that "making public sector agencies actually use the data internally for evidence-based policy-making in order to defend the spend has caught on like wildfire".[37] Within private sector partners that work with government agencies, such as Socrata[38] and ESRI,[39] there has been a sizable shift from simply managing open data platforms to building capacity to visualise and manage real-time government performance data. Whether or not performance management will take hold across countries and levels of government, there is a growing demand for products and platforms that enable data-driven governance.

Many pundits have opined that data is the new oil.[40] The value and strategic importance of data in modern government suggests a critical role for the open data movement. In many ways, making government information more transparent and available transcends arguments about the right success markers, the roles of champions, adoption rates, or attentiveness to user demand. Truly opening up public data holds the promise of creating a more level playing field, ensuring that early movers and big players do not dominate the use of public data sources. As government becomes a more sophisticated and smarter user of data in establishing and pursuing its goals, open data may well form the bridge that connects data-informed policy objectives to citizen engagement.

Conclusion

This chapter opened with the shuttering of open.whitehouse.gov, signalling both the political risks that face the open data movement in Australia, Canada, New Zealand, and the US and just how robust and dedicated the open data community has become in these countries. The stories of progress and the remaining gaps to be addressed are not intended to represent all that has happened in open data in recent years, but rather a summary of the most important trends that practitioners and advocates have identified as having an impact on their work.

Based on the challenges identified and best practices that have emerged, policy-makers should focus on three distinct areas of reform. First, foster more standards around the collection and publishing of open data, particularly at the local level. Second, push government staff and departments to share and make use of each other's data, in particular, by investing in tools that enable them to share and a culture that encourages public servants to leverage each other's data when drafting policy and improving operations. Third, publish open data in response to demand and work to identify the kinds of public problems for which the use of open data can help create innovative solutions.

The stories of progress identified here demonstrate that the hard-fought victories to increase the availability of open data have created real value for the governments, citizens, and businesses of North America, Australia, and New Zealand. On the other hand, the persistent challenges facing the advance of open data in these countries should also serve as a warning. For open data to move beyond early successes and quick wins, governments must address the uneven demand for open data, better identify the users and use cases of open data, and push for long-term local adoption. The state of open data in these nations is currently neither a crisis nor a celebration, but these steps will be critical to ensuring that open data becomes a truly integral, sustainable, and effective pillar of 21st century governance.

Further reading

Civic Analytics Network. (2017). An open letter to the open data community. *Data-Smart City Solutions*, 3 March. https://datasmart.ash.harvard.edu/news/article/an-open-letter-to-the-open-data-community-988

Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., & Almasi Doshi, E. (2013). Open data: Unlocking innovation and performance with liquid information. *McKinsey Digital*. https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information

Tauberer, J. (2014). *Open government data: The book* (2nd edition). https://opengovdata.io/

About the authors

David Eaves is a public policy entrepreneur and expert in information technology and government. He has worked closely with municipal and federal leaders in Canada and the United States on the intersection of technology, open data, and governance. He is currently the Faculty Director of digital HKS at the Harvard Kennedy School of Government. You can follow David at http://www.twitter.com/daeaves and learn more about his work at https://eaves.ca.

Ben McGuire is a public policy master's candidate at the Harvard Kennedy School. He previously worked in strategy research in the education sector, as well as in organising and data analytics in American political campaigns. You can follow Ben at http://www.twitter.com/bean_mcguire.

Audrey Carson has contributed to technology and public policy research at the Harvard Kennedy School, the Berkman-Klein Center at Harvard Law School, and the MIT Media Lab. She currently works as a strategic communications consultant advising companies and organisations on external communications, corporate positioning, and public affairs.

How to cite this chapter

Eaves, D., McGuire, B., & Carson, A. (2019). Open data around the world: North America, Australia, and New Zealand. In T. Davies, S. Walker, M. Rubinstein, & F. Perini (Eds.), *The state of open data: Histories and horizons.* (pp. 517–534). Cape Town and Ottawa: African Minds and International Development Research Centre. http://stateofopendata.od4d.net

Key informant interviews

In the process of researching this chapter, interviews were conducted with the following key informants:

Diego Cuesy, Former Open City Coordinator at Laboratorio para la Ciudad in Mexico City; Jay Daley, Managing Director of techobscura Ltd and former CEO of the .NZ registry; Valentina Delgado, Open Data Strategy Consultant at Laboratorio para la Ciudad in Mexico City; Jason M. Hare, Open Data Principal at Jason M. Hare Associates; Alannah Hilt, Business Lead, Open Government Portal, Treasury Board of Canada Secretariat; Tyler Kleykamp, Chief Data Officer for the State of Connecticut, United States; Scott McQuarrie, Geomatics Coordinator for the Township of Langley, Canada; David Moore, Executive Director of the Participatory Politics Foundation and Co-founder of Sludge; Andrew Nicklin, Director of Data Practices at the Center for Government Excellence at Johns Hopkins University; Denice Ross, Phaze Zero Project at New America; Gabe Sawhney, Executive Director of Code for Canada; Paul Stone, Open Government Data Programme Leader for the Secretariat for Open Government Data and Information Programme, New Zealand; David Wasylciw, Founder of OpenNWT.

Endnotes

1 Miller, Z.J. (2017). The White House will keep its visitor logs secret. *Time*, 14 April. http://time.com/4740499/white-house-visitor-logs-public-record-trump/

2 Tarantola, A. (2017). Trump administration is killing its open data portal. *Engadget*, 14 April. https://www.engadget.com/2017/04/14/trump-admin-killing-open-data-portal/

3 President's Management Council and Executive Office of the President. (2018). *President's management agenda*. https://www.whitehouse.gov/wp-content/uploads/2018/03/Presidents-Management-Agenda.pdf

4 Knell, N., Pittman, E., Towns, S., & Mulholland, J. (2014). Open data policies in state and local government. *GovTech,* 17 March. http://www.govtech.com/data/Are-Governments-Committed-to-Open-Data-Interactive-Map.html

5 https://www.anz.co.nz/about-us/economic-markets-research/truckometer/

6 Byrne, D.P., Nah, J.S., & Xue, P. (2018). Australia has the world's best petrol price data: FuelWatch and FuelCheck. *Australian Economic Review*, *51(4)*, 564–577. https://doi.org/10.1111/1467-8462.12302

7 https://www.remix.com/

8 https://www.tylertech.com/resources/blog-articles/connecticut-data-sparks-winning-app-for-opioid-crisis

9 Riddell, D.J. (2014). From the ground up: The story of the Canadian boreal forest agreement. https://www.researchgate.net/publication/303328044_From_the_Ground_Up_The_Story_of_the_Canadian_Boreal_Forest_Agreement

10 https://www.policedatainitiative.org/

11 http://data.torontopolice.on.ca /

12 http://www.policedata.nz/

13 https://elections.bc.ca/resources/voting-results/provincial-by-elections-results/

14 https://www.digital.nyc/startups/heat-seek-nyc

15 https://www.govtrack.us/

16 http://www.councilmatic.org/

17 Key informant interview.

18 Web Foundation. (2017). *Open Data Barometer – Global report*. 4th edition. Washington, DC: World Wide Web Foundation. https://opendatabarometer.org/4thedition/report/

19 https://opendatabarometer.org/4thedition/methodology/

20 https://www.linz.govt.nz/data/linz-data-service

21 Madrigal, A. (2009). Data.Gov launches to mixed reviews. *Wired*, 21 May. https://www.wired.com/2009/05/datagov-launches-to-mixed-reviews/

22 Key informant interview.

23 Key informant interview.

24 Key informant interview.

25 Key informant interview.

26 https://publicsectordigest.com/open-cities-index-results-2017

27 https://opendatabarometer.org/

28 Key informant interview.

29 https://communitydata.ca/SaultSteMarieConsortium

30 Key informant interview.

31 Key informant interview.

32 Key informant interview.

33 Key informant interview.

34 Key informant interview.

35 Goldsmith, S. (2015). Chicago's data-powered recipe for food safety. *Data-Smart City Solutions*, 21 May. https://datasmart.ash.harvard.edu/news/article/chicagos-data-powered-recipe-for-food-safety-688

36 Howard, A. (2012). Predictive data analytics is saving lives and taxpayer dollars in New York City. *O'Reilly Media*, 26 June. https://www.oreilly.com/ideas/predictive-data-analytics-big-data-nyc

37 Key informant interview.

38 https://socrata.com/

39 https://www.esri.com/en-us/about/about-esri/overview

40 Anon. (2017). The world's most valuable resource is no longer oil, but data. *The Economist*, 6 May. https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data