

# **Open Data in Scotland**

**A Blueprint for Unlocking Innovation,  
Collaboration and Impact**

**August 2024**

# **Open Data in Scotland**

## **A Blueprint for Unlocking Innovation, Collaboration and Impact**

Jack Lord, Open Data Services Co-operative Ltd.

### **Acknowledgements**

This report was written by Jack Lord from Open Data Services Co-operative Ltd and funded by the Data Division within the Digital Directorate of the Scottish Government. The content of this report does not reflect the official opinion of the Scottish Government. Responsibility for the information and views expressed therein lies entirely with the authors. The author would like to thank both those who ran and those who took part in workshops and interviews for their time and contributions.

### **Executive Summary**

The Scottish Government has the opportunity to unlock the potential of open data to drive innovation, foster collaboration and generate impact. To do this, it needs to set out a strong vision for open data in Scotland, then coordinate and promote its publication and use.

Scotland can look to a global open data movement for inspiration. Developments over the last decade have provided a solid foundation for best practices and technical approaches, which can make data publication simpler and more sustainable. The result is higher quality data that can be analysed and repurposed in multiple contexts, which shortens the route to impact. Worldwide, governments are adopting an intentional approach to data publication, which reframes open data as a useful tool to achieve policy goals and societal impacts, rather than as a technical output aimed at a niche audience. Scotland is progressing on both fronts but more can be done to unlock the full social and economic potential of open data.

The first step is to establish a vision for open data in Scotland, which positions open data as a key tool for advancing policy priorities, rather than as a tick box exercise. Crucially, this approach requires collaboration among stakeholders: from data publishers, to the public and private sector, and civil society. In turn, this collaborative approach can forge new partnerships, drive data publication that is driven by social and economic needs, and foster a long-term commitment to open data that can achieve lasting impact.

The second step is to develop a pathway for implementation that will lead to impact; by making informed decisions about the technologies and standards used to publish and access data; and by providing appropriate support to publishers. This approach relies on transparent reporting and continuous improvement. By measuring impact and celebrating success, the Scottish Government can maximise the benefits of open data in Scotland, in line with the commitments to improving open data provision in the Open Government action plan<sup>1</sup>.

The full report provides an overview of trends in open data, and reports on feedback from open data practitioners in Scotland. We set out five recommendations, which outline how the Scottish Government can position itself as a leader in open data and evidence based policy making.

---

<sup>1</sup> [2024-25 Scottish Budget - Pre-budget Scrutiny Report](#) (2023)

# Table of contents

Acknowledgements	1
Executive Summary	1
Table of contents	3
1. Introduction	4
2. Open data: literature review	5
2.1 Defining open data	5
2.2 Trends in open data	6
2.3 The value and impact of open data	9
2.4 Producing, using and achieving impact with open data	13
2.5 What do we mean by 'guidance' for open data	17
3. Open data in Scotland	25
3.1 Policy environment	25
3.2 Publication landscape	27
3.3 Findings from workshops	28
3.4 Opportunities for the future	30
4. Recommendations	32
4.1 Set out a vision that makes the case for open data	32
4.2 Embrace long-term collaboration and commitment	35
4.3 Create an implementation strategy that is intentional, realistic, transparently monitored, and celebrates success	36
4.4. Set the standards for open data publication	38
4.5 Provide targeted technical guidance and support for publishers and users	40
5. Summary and next steps	43

# 1. Introduction

This report was commissioned by the Open Data Team in the Scottish Government Digital Directorate to conduct research on the domestic and international open data landscape. The purpose of this research was to understand the current status and requirements in Scotland, and the global context in which they sit. These findings will inform us in bringing existing guidance for open data up to date, and identify any new guidance that needs to be provided.

The research contained two key strands: a series of public workshops (run by Fractals Co-operative) and desk research on the state of open data. In response to the research findings the scope of this report was broadened to cover guidance in the more expansive sense, including technical and policy elements.

The report is structured in three parts. Part one summarises the current state of open data globally, and with reference to Scotland. It examines some of the major themes and best practices that have emerged with respect to publishing data that is valuable, impactful and sustainable. Part two explores the Scottish context in more detail, including brief summaries of the major findings from the public workshops. Part three discusses some particular points identified in the research and makes five recommendations to tackle these and move open data in Scotland forward.

## 2. Open data: literature review

### 2.1 Defining open data

There is relative unanimity about what 'open data' and 'open government data' mean. Open Knowledge Foundation's widely used Open Definition<sup>2</sup> states that open data is:

"[Data that] anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)"

Open government data is any data produced, collated or commissioned by public bodies, which conforms to this definition or an equivalent open definition.<sup>3</sup>

This simple definition can be expanded upon in two main directions. One approach is **output-based**, using the legal and technical characteristics of data to define its openness. The full Open Definition takes this approach, requiring open data to be:

- Published under an open license<sup>4</sup>
- Accessible for free or at cost
- In a machine-readable format
- In an open format

The other approach is **process-based**, and judges the openness of data on the principles, policies and processes that shape data governance and publication. Examples of this contextual approach include the original 2007 Open Government Data Principles<sup>5</sup>, which included requirements that data be timely, complete and granular, and the wider concept of Open ICT for Development, which stressed universal access and participation and collaborative production methods.<sup>6</sup>

---

<sup>2</sup> Open Knowledge Foundation, [Open Definition 2.1](#) (n.d.).

<sup>3</sup> Ubaldi, Barbara. "[Open government data: Towards empirical analysis of open government data initiatives](#)." (2013).

<sup>4</sup> An open license is one which grants a minimum set of permissions and imposes only certain 'acceptable' conditions on use. Open Knowledge Foundation, [Open Definition 2.1](#) (n.d.).

<sup>5</sup> [Open Government Data Principles](#), (2007).

<sup>6</sup> Davies, T. [Open ICT for Development](#), (n.d), [Open Data as Strategy](#), (2016).

The output-based definition of open data is now dominant but questions over process and context remain relevant, even if they are now approached via a data governance and strategic lens (see sections 2.2 and 2.3).

## 2.2 Trends in open data

The term open data was coined in 1995, in reference to the free exchange of scientific data. The Open Government Data Principles were written in 1997, with the US adopting a transparency and open government directive in 2009, and the first International Open Data Conference taking place in 2010. Open data has since become increasingly important to policy initiatives across the world, as a transparency mechanism, a practical tool to tackle policy issues, and as a means to measure and monitor progress on other goals.<sup>7</sup>

Underneath this narrative however the priorities and methodologies of the open data movement have changed significantly. Verhulst et al. describe several chronological 'waves' of open government data that act as a useful framework<sup>8</sup>:

- **First Wave:** data is released piecemeal in response to requests based on freedom of information laws.
- **Second Wave:** data is released using an "open by default" principle, based on the idea that open data will catalyse innovation, evidence-based decision-making, civic engagement, and a culture of collaboration. Limited attention is paid to demand during this phase. Scotland currently has "open by default" as a principle of its own open data strategy.
- **Third Wave** (current phase): a purpose-driven approach supplants the sometimes scattershot publication of the Second Wave. This prioritises impactful reuse and broader collaboration to produce complex datasets needed to address complex challenges.
- **Fourth Wave** (predicted): generative AI makes it easier to use and analyse data, improve data quality, and allow the creation of synthetic datasets where privacy or availability is a concern.<sup>9</sup>

---

<sup>7</sup> Davies, T. et al, (Eds.). [The State of Open Data: Histories and Horizons](#) (2019)

<sup>8</sup> Verhulst, Stefaan, et al. "The emergence of a third wave of open data: How to accelerate the re-use of data for public interest purposes while ensuring data rights and community flourishing." *Available at SSRN 3937638* (2020).

<sup>9</sup> The Governance Lab, [Toward a Fourth Wave of Open Data?](#) (2023)

One way to interpret the ‘waves’ of open government data is as an evolution of the motivation for publication from:

- A **compliance** approach, driven by reactive responses to external stakeholders.
- A **principles** approach in which openness is integrated into the data governance principles used by public bodies.<sup>10</sup>
- An **intentional** approach, in which open data publication is planned and targeted to align with other policy priorities and to fit the resource and capability constraints faced by publishers.

All three of these attitudes can be found among open data stakeholders in Scotland.

The transition from the second to third wave in the public sector has been and continues to be driven by:

- An increased public awareness and literacy around data availability and use;
- Improved understanding about why and how data practitioners can be embedded in the broader processes and governance structures of local communities directly affected by open data projects; and,
- A targeted, demand-driven approach as a more practical and sustainable option for organisations with limited financial and technical resources.

The following case studies are examples of the third wave approach to open data.

### **Case study 1: [Scotland’s financial transparency Open Government Partnership \(OGP\) commitments](#)**

#### **Purpose**

- Improve reliability and accessibility of fiscal data, so that a wide range of users can understand and reuse it.

---

<sup>10</sup> Data governance is defined by DAMA as: “The exercise of authority, control and shared decision making (planning, monitoring and enforcement) over the management of data assets”. Data governance is therefore important because it defines why and how open data gets published, what the outputs look like, and who is involved in the process. DAMA. Earley, S., et al (Eds.), The DAMA Guide to the Data Management Body of Knowledge (DAMA-DM BOK). (2017).



## Partnerships

- Extensive collaboration within and between departments, civil society and domain experts.

## Approach

- Identification of high value and high impact datasets. Making long-term changes to improve access via stakeholder engagement and technical improvements.

## Impact

- Procurement transparency with data published in line with the Open Contracting Data Standard (OCDS).
- Standardisation and visualisations of extensive disclosures on infrastructure investment.
- Ongoing work on improved transparency around the Scottish Budget.

## Case study 2: [International Aid Transparency Initiative \(IATI\)](#)

### Purpose

- Meet transparency commitments set out in the [Accra Agenda](#) to help a diverse range of stakeholders improve the effectiveness and efficiency of aid and development finance.

### Partnerships

- Providers of development assistance, partner countries, multilateral institutions, private sector and civil society organisations and others.

### Approach

- Shift from general disclosure to the targeting of data disclosures and data use initiatives for specific purposes.

### Impact

- Interactive map of the African Development Bank's investments in Africa helps with allocation of resources.
- Aligning international aid flows with national priorities improves government decision-making in Liberia.
- Sectoral initiatives to understand where aid is going and how effective it is (eg. [Pacific Aid Map](#), [DFI investments in gender](#), [global air quality funding](#)).

While the Scottish Government is shifting to a more intentional approach to produce complex and connected datasets, civil society is largely driven by a principles approach in which data should be open by default, and many smaller public sector organisations see open data as a compliance issue. This diversity of approaches is almost inevitable but it can lead to misunderstandings and uncertainty if unaddressed.

## 2.3 The value and impact of open data

Value and impact are ultimately why public sector organisations publish open data. Understanding and documenting the expected value and impact of open data can help build support for open data, feed into choices over what and when to publish, and measure outcomes. This section describes some of the most important ways to think about the impact and value of open data and outlines the most common techniques for measuring that impact and value.

### 2.3.1 Impact

The expected impact of open data can take many different and overlapping forms.

- **Transparency, accountability and empowerment.** Open data can support transparency and accountability when it is used as a public record of actions taken by public actors and institutions. Open data improves accountability because it allows the public to engage with and understand such disclosures (and for public bodies to respond).<sup>11</sup> Open data can help citizens to be more engaged and empowered by allowing direct access to information about local communities and areas of interest, and by creating the infrastructure for accessible information services aimed at non-technical audiences. Examples include:
  - Open budget, spending and procurement data have made complex public sector finances more accessible and comprehensible.
  - Registers of political interests and information on elections and parliamentary activity have allowed citizens to engage more closely with democratic processes.
  - Open data about local services, facilities and socio-economic indicators has improved the ability of citizens to understand and engage with local issues.

---

<sup>11</sup> Peixoto defines public transparency as ‘the disclosure of actions taken by public actors and institutions’. Peixoto, Tiago. (2013). The Uncertain Relationship Between Open Data and Accountability. UCLA law review. University of California, Los Angeles. School of Law. 60. 200-213.

- **Innovation and economic growth.** Open data can contribute to innovation and economic growth, primarily through datasets that directly form part of value-added activities where data is the primary product or service. These include enhancing, interpreting or combining data. Data sets may also indirectly support value-added activities where data is not the primary product or service. For example, these allow innovations in processes, improving market function, providing efficiency gains or reducing risk. Examples include:
  - Open transport datasets led to passenger information and mapping services.<sup>12</sup>
  - Open company and financial information led to due diligence and compliance services.<sup>13</sup>
  - Open satellite mapping increased the rate of gold discoveries and the share of the market held by new entrants.<sup>14</sup>
  - Open data on public and publicly-funded services has made it easier for end-users and intermediaries to find suitable services.<sup>15</sup>
  
- **Management and monitoring.** Open data can support the management and monitoring of public policy goals by providing indicators of progress derived from existing datasets, and by identifying and filling gaps in current data collection and generation. Examples include:
  - Scotland's National Performance Framework relies on multiple open data sets to provide visibility on progress towards agreed outcomes.<sup>16</sup>
  - Shortcomings in available gender data to monitor progress towards the Sustainable Development Goals (SDGs) has led to a movement to improve funding for gender-based data systems.<sup>17</sup>
  
- **Collaboration and coordination.** Open data can improve collaboration and coordination around public policy goals by enabling data to be shared across government with a minimum of friction, and allowing data to be shared and collated across the public, private and third sectors. Examples include:
  - Publication on financial commitments, transactions and results of international development and humanitarian activities has allowed providers and recipients of assistance to better plan, evaluate and coordinate their work.<sup>18</sup>

---

<sup>12</sup> Deloitte, [Assessing the value of TfL's open data and digital partnerships](#), (2017)

<sup>13</sup> Lord, J. (2019) [Open Data and Corporate Ownership](#). In T. Davies, S. Walker, M. Rubinstein, & F. Perini (Eds.), *The State of Open Data: Histories and Horizons*. Cape Town and Ottawa: African Minds and International Development Research Centre.

<sup>14</sup> Nagaraj, A. '[The Private Impact of Public Data: Landsat Satellite Maps Increased Gold Discoveries and Encouraged Entry](#)', *Management Science* 2022 68:1 (2022)

<sup>15</sup> Open Data Institute, [The Role of Data in Unlocking the Potential of Social Prescribing](#)

<sup>16</sup> [National Performance Framework](#)

<sup>17</sup> <https://data2x.org/state-of-gender-data/financing/>

<sup>18</sup> [International Aid Transparency Initiative. Case Studies](#)

- Sharing and publication of data on climate and natural disaster risks allows governments, scientists, local communities and private sector stakeholders to better plan, maintain and protect infrastructure and physical safety.<sup>19</sup>

### 2.3.2 Value

Similarly, there are many different models for thinking about the value of open data. Maximising the value of open data means realising multiple types of value, rather than pigeon-holing data into one particular area. Examples include:

- **Social and economic value.** Some data is designed for explicitly social ends (for example, data on democracy or about local communities), while other data is published primarily for its potential impact on the economy. For example, mapping and transport data. Over time new use cases will tend to blur these distinctions. There is an ecosystem of due diligence and analytical companies built around open company data, for example, but this data is also widely used by civil society organisations working on integrity or environmental issues. Similarly, 360Giving, which allows open data on grant funding, realises value across this spectrum: providing transparency on the equitability of the sector; helping grant-makers to collaborate to be more effective and intentional; and, making the process of finding grant funding more efficient.<sup>20</sup>
- **Planned and emergent value.** For some open data the value of publication might be understood up front, or clarified during the research and design phase, as part of the kinds of intentional and planned publication associated with the third wave of open data. This is particularly relevant for complex and collaborative datasets, where the investment needs to be justified and the outputs need to meet specific and well understood user needs.<sup>21</sup> But the value of open data can also emerge over time, as new use cases are developed. In this case it may be preferable to reduce the time spent on planning and design and focus instead on making data accessible, discoverable and easy to use.<sup>22</sup>
- **Primary and secondary value.** The most obvious type of value stems from direct or indirect use of open data, as new activities and services are enabled or

---

<sup>19</sup> Open Data Services, [How open data standards can help mitigate climate and disaster risk](#), (2023).

<sup>20</sup> <https://theodi.org/insights/reports/measuring-the-impact-of-data-institutions-report/>

<sup>21</sup> <https://os4d.opendataservices.coop/development/research/>

<sup>22</sup> Robinson, David G. and Yu, Harlan and Zeller, William P. and Felten, Edward W., Government Data and the Invisible Hand (Fall 2009). Yale Journal of Law & Technology, Vol. 11, p. 160, (2009) Available at SSRN: <https://ssrn.com/abstract=1138083>

improved. But the process of producing open data, and the culture it encourages, can also create positive spillover effects, like improvements to data governance and IT processes, or the creation of new working relationships within and outside the public sector.<sup>23</sup> These types of a secondary value can be a conscious by-product of ‘open data as strategy’, in which the approach required to produce useful open data is valued on a par with the technical outputs.<sup>24</sup>

- **Wide value and narrow value.** Another difference exists in the scale and breadth of value that an open dataset delivers (or is expected to deliver). The concept of ‘high value datasets’ used in the EU Open Data Directive, for example, is based on open data having significant impacts for large numbers of beneficiaries.<sup>25</sup> But significant value can also emerge from ‘highly valued datasets’ that are of vital interest to a smaller audience.<sup>26</sup>
- **Output and outcome value.** Finally, value can be seen in terms of outputs (for example, the number of datasets produced and the technical quality of those datasets) and also in terms of outcomes (what the underlying information allows direct or indirect users to achieve). Large quantities of well-ordered, high-quality data is the ideal situation but, nonetheless, significant value can be derived from poor data of dubious quality when users embrace the ‘messiness’ of such information. For example, the Organised Crime and Corruption Reporting Project’s Aleph software is designed to help journalists and investigators navigate the complexity of official and unofficial records of corporate activity.<sup>27</sup>

### 2.3.3 Measurement

Measuring the value and impact of open data is important to build and retain support for publication and to make sure that plans are informed by evidence about what works. That said, measurement is recognised as difficult and detailed evidence on impact remains scarce.<sup>28</sup> The main ways in which value and impact can be measured are:

---

<sup>23</sup> Goëta, S. and Davies, T. (2016) “The Daily Shaping of State Transparency: Standards, Machine-Readability and the Configuration of Open Government Data Policies”, *Science & Technology Studies*, 29(4), pp. 10–30. doi: 10.23987/sts.60221.

<sup>24</sup> Davies, T, [Open Data as Strategy](#), (2016).

<sup>25</sup> [Open Data Directive](#)

<sup>26</sup> Davies, T, [High value datasets: an exploration](#), (2019)

<sup>27</sup> OCCRP, [Aleph](#)

<sup>28</sup> Lämmerhirt, D. & Brandusescu, A. (2019) Issues in Open Data - Measurement. In T. Davies, S. Walker, M. Rubinstein, & F. Perini (Eds.), *The State of Open Data: Histories and Horizons*. Ahmadi Zeleti, Fatemeh, Rhoda Kerins, and Helena Campbell. "[Towards Open Data Impact Evaluation Framework—An Empirical Analysis of the Demand-side Response](#)." (2023).

- **Qualitative case studies** of particular data initiatives or areas of impact. These can be particularly useful in making the value of open data clear to non-technical audiences and in identifying priority areas for improvement. Case studies are vulnerable to cherry-picking, however, and so are likely to work best when produced by sector experts and end users who can provide honest feedback.
- **Quantitative case studies** that use economic modelling or other methods to estimate the benefits of open data.<sup>29</sup> These can be particularly useful to overcome scepticism over potentially complex and costly publication and to make cost-benefit analysis before and after publication.<sup>30</sup> It can be difficult to isolate the impact of open data in economic modelling, however, and there is no one-size-fits-all methodology so this can be a complex approach.
- **Quantitative indicators** that demonstrate how open data is being used can be useful to understand impact over time and to integrate open data initiatives into existing Monitoring, Evaluation and Learning frameworks. While potentially useful, such indicators can be difficult to design, and there may be a temptation to choose unambitious targets if they are used to measure success or failure.<sup>31</sup>
- **Surveys and catalogues** of data use and impact among stakeholders can be useful to understand the uptake and usage of open data inside and outside the public sector. The GovLab has collated a collection of global impact case studies and Ireland uses a regular survey to gauge the use of open data.<sup>32</sup>

## 2.4 Producing, using and achieving impact with open data

The optimal practical approach to producing open data has also changed as the field has matured because publication is now recognised as a **socio-technical** problem—that is a problem with ‘organisational, human, material, and technological’ dimensions.<sup>33</sup> This means that, while sound technical implementation remains fundamental, the

---

<sup>29</sup> [Creating Value through Open Data: Study on the Impact of Re-use of Public Data Resources](#)

<sup>30</sup> The UK Government, for example, quantified the cost to businesses of [collecting beneficial ownership information](#), which was published as open data.

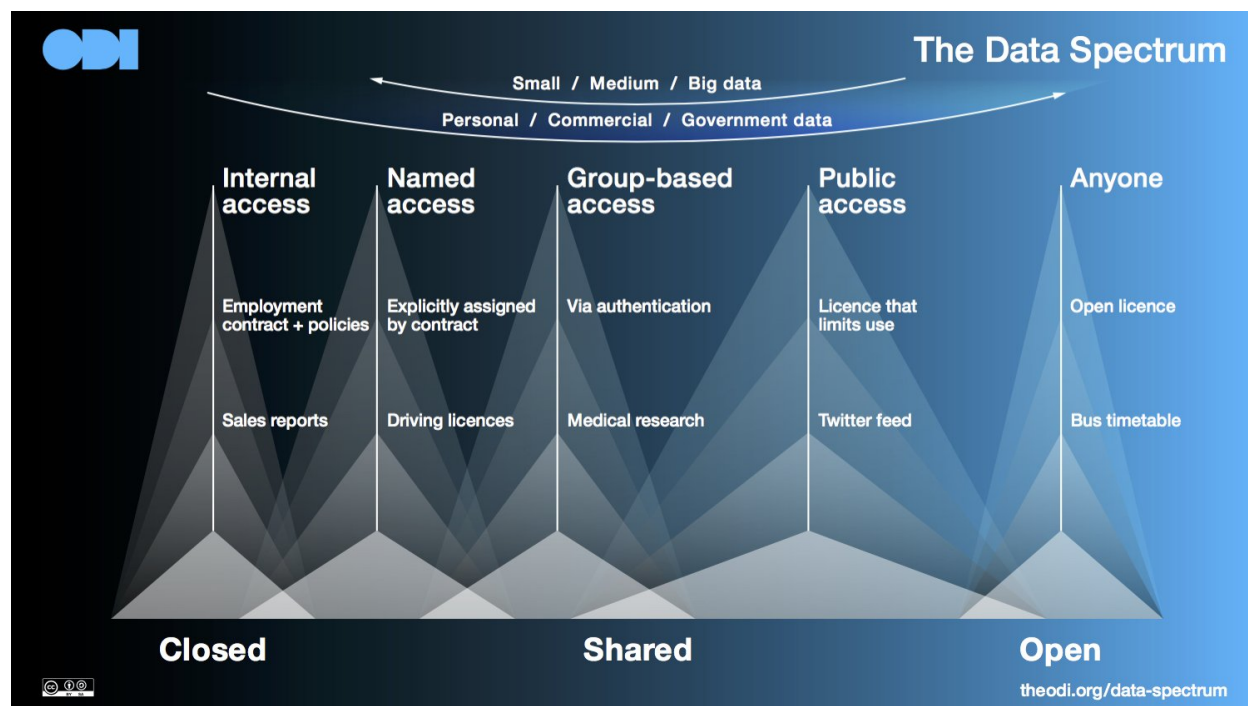
<sup>31</sup> Ahmadi Zeleti et al, [Towards Open Data Impact Evaluation Framework – An Empirical Analysis of the Demand-side Response](#), Proceedings of the 56th Hawaii International Conference on System Sciences, (2023)

<sup>32</sup> [GovLab, Open Data's Impact](#)

<sup>33</sup> Dawes, SS, Vidasova, L, Parkhimovich, O. (2016). Planning and designing open government data programs: an ecosystem approach, Government Information Quarterly, <http://dx.doi.org/10.1016/j.giq.2016.01.003>

environment in which open data is published is seen as a key limiting and enabling factor.

**Figure 1: The Data Spectrum, from closed to open.**



Source: the Open Data Institute, under a [CC-BY](https://creativecommons.org/licenses/by/4.0/) license

Rather than a narrow focus on purely open data, we can derive benefits from sharing data in different ways (see figure 1, above), which range from granting limited access, (such as, by restricting the audience or redacting parts of the dataset) to full access for reuse under an open licence. Sharing and opening data can be mutually reinforcing, as many of the organisational, technical, and policy barriers are the same.

This section discusses the idea of data value chains (which incorporate the full complexity of open data publication, broadening the discussion to beyond the lifecycle of just the data itself), some of the most important data governance principles for open data, and the implications of this for technical guidance about open data.

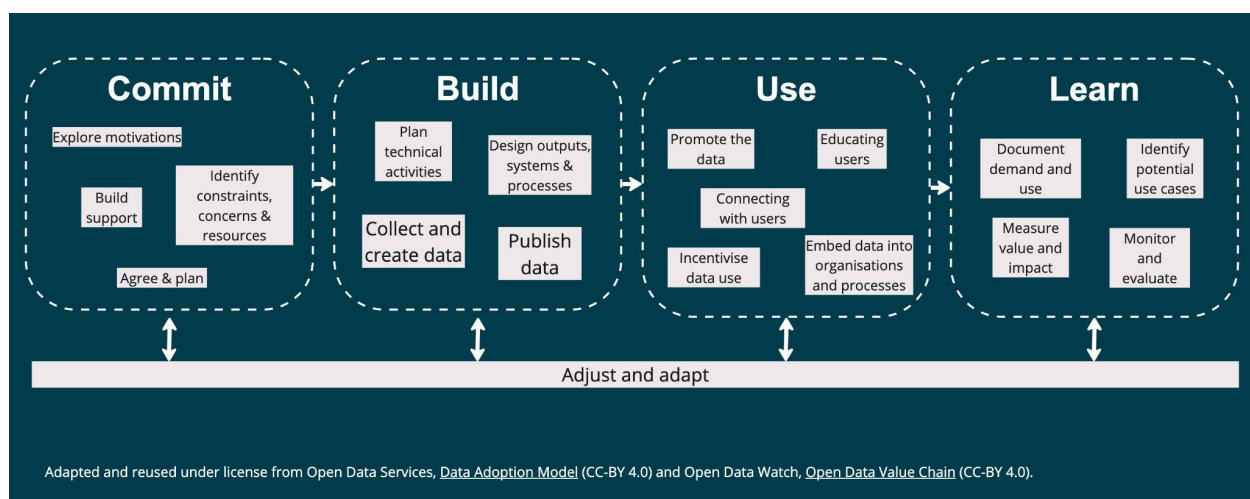


## 2.4.1 Data value chains

An open data value chain is the set of activities that lead to the identification, publication and use of open data, designed to maximise the impact of open data at each stage in the open data lifecycle. An open data value chain can be a design tool, allowing the right activities to be identified to maximise impact, and a diagnostic tool, allowing points of friction and areas for improvement to be identified.<sup>34</sup>

Figure 2 below shows an example open data value chain with some non-exhaustive example activities at each of the four main stages.

**Figure 2: An example open data value chain**



Source: Adapted and reused under license from Open Data Services, [Data Adoption Model](#) (CC-BY 4.0) and Open Data Watch, [Open Data Value Chain](#) (CC-BY 4.0)

An open data value chain is a lens through which to view the full lifecycle of open data, focusing on the impact - potential, or realised - of each stage.

- **Commit:** This stage involves getting to an agreement to publish. This may involve implementing an existing policy or directive, or building support for a new initiative. This stage also incorporates important design elements, such as defining the purpose of publication, setting data governance principles and policies, identifying key stakeholders and building partnerships, and understanding the likely material, organisational or human constraints.

<sup>34</sup> Data2X/Open Data Watch, [The Data Value Chain: Moving from Production to Impact](#)



- **Build:** This stage involves the activities that lead to data publication itself, including data collection and mapping, technical design and negotiation, designing and implementing data governance practices, creating data pipelines, and publishing and cataloguing the data.
- **Use:** This stage involves turning published data into a resource that is actively used and understood. Relevant activities may include data analysis itself, promoting data to potential users, forging partnerships, and capacity-building for potential users.
- **Learn:** This stage involves understanding how data is being used, the impact it has and any areas where potential is not being fulfilled, or where implementation issues are holding back impact.

There are three overarching concepts that apply to all open data value chains:

- **Iteration:** A data value chain is not linear and there is no definitive endpoint. Instead, adaptation and adjustment should take place in parallel and learning from each stage should feed back into other stages.
- **Collaboration:** Open data value chains are collaborative by nature. They bring together technical and non-technical stakeholders and stakeholders from across the public sector with civil society and the private sector. Stakeholders' interests and levels of involvement will vary across the open data value chain. A well-functioning open data value chain will result in coordinating the needs and challenges of these different groups.
- **Context:** The design of an open data value chain is highly context-dependent. This context covers multiple levels, from the high-level of the legal, regulatory and policy environment, to the constraints and priorities of implementing organisations, and down to the skills and interests of teams and individuals. There is no off-the-shelf solution for open data publication and an appropriate open data value chain will be tailored to specific needs.

Finally, open data value chains are useful for highlighting the complexity of data publication and therefore the many ways that expected impact can fail to materialise if implementation is not done well.<sup>35</sup> Critical reflection on activities across the full open

---

<sup>35</sup> GovLab identified five major areas that need attention in order to maximise the impact of open data: problem and demand definition, capacity and culture, partnerships, risks, and governance. GovLab, [Periodic Table of Open Data's Impact Factors](#).

data value chain can be a useful diagnostic tool in identifying and fixing issues with key activities.

## 2.5 What do we mean by ‘guidance’ for open data

This research was originally designed to provide advice on ‘technical guidance’ for open data. As the previous sections have shown, however, open data publication is much more than just a technical problem and a wider concept of guidance is required.

The core purpose of open data guidance should be to help publishers design and navigate the technical and non-technical activities in a data value chain, reducing friction and identifying key touch points that can maximise positive impact.

This section discusses three layers of guidance that can achieve these goals: creating an enabling environment for open data; setting the high level direction for open data publication; and providing tools, resources and support. In Section 4, we discuss which of these are practical and realistic for the Scottish Government to provide.

### 2.5.1 An enabling environment

The impact of detailed guidance on open data publication will depend on the quality of the enabling environment in which they sit. Three elements of that enabling environment, which policy initiatives can help improve, are particularly relevant:

- **Direction and leadership.** Strong direction and ongoing leadership on the importance of openness and open data can help motivate publication for public sector organisations, set expectations on what will be published, and reduce the perceived risks of transparency and data publication. Detailed open data policies (or, in certain circumstances, legal and regulatory requirements) can further reduce uncertainty, allowing organisations to then concentrate on implementation. While the Scottish Government has numerous commitments to open data the perception is that the ambition is relatively low and implementation lags behind, in part due to misaligned or inappropriate expectations, and a lack of high-level leadership in the area.
- **Culture and trust.** A culture aligned with the principles of openness and collaborative working is important, particularly when considering how to move beyond a risk-averse, compliance mindset. An open culture built on trust can help with several key challenges: creating genuine enthusiasm around the

benefits of transparency, building cross-sector partnerships designing initiatives with the input and feedback of stakeholders, and seeing open data publication as a collaborative and collective commitment in which the public sector, the private sector and civil society need to work together. While the Scottish Government is committed to participatory and open principles there is a perception that these are not universally held across the public sector and the relationship between data publishers and civil society can sometimes feel combative rather than collaborative.

- **Capability and resources.** Closing the gap between ambition and implementation relies on public sector organisations having the capability to publish data, and end users having the capability to analyse, understand and deploy data. The distributed responsibility for publication in Scotland (see Section 3.2) makes this a difficult problem as organisations often have very limited technical resources and are starting from scratch on open data. These same problems also apply to end users, who may be interested in the information that has been published, but lack the skills or time to process technical outputs. These deficits may be tackled via a general and long-term approach to improving digital skills and data literacy, the provision of some central technical resource, or by encouraging ‘infomediaries’ to contextualise and explain open data to non-technical audiences.<sup>36</sup>

## 2.5.2 High level direction

Setting the high level direction via a common data governance framework for publishing open data can help remove uncertainty about what data needs to be published, how that data should be published and managed, and the quality and technical standards that are expected. This reduces the planning and decision burden on data publishers, makes the data more predictable and usable for end users, and creates shared expectations on data availability, quality and other factors.

Some particularly relevant aspects of data governance for open data in Scotland include: policies, strategies, principles, quality standards, technical standards, discoverability and availability.

**Open data policies** are a common tool for establishing the ground rules for publication and can create a solid foundation for impactful data publication and use. The most

---

<sup>36</sup> Peter A. Johnson, Sarah Greene (2017). Who Are Government Open Data Infomediaries? A Preliminary Scan and Classification of Open Data Users and Products. UWSpace. <http://hdl.handle.net/10012/16730>

important elements of such a policy include setting an overarching purpose, specifying a justification for publication (including via a legal framework), establishing operational approaches, specifying quality and technical standards, assigning responsibility for key tasks, providing a framework for producer-user relationships, and discovering and stimulating demand and supply.<sup>37</sup>

A **strategy** can be useful for operationalising a policy, avoiding the risk that a policy remains a paper exercise only. South Korea, for example, established an open data law in 2013, backed by regular high-level plans that encourage a purpose-directed approach to publication and a more granular plan that directs data releases.<sup>38</sup> The Scottish Government has a number of policy initiatives related to open data but has been less successful in establishing shared expectations on the practicalities among publishers, leading to large divergences in practice and a publication landscape that is generally fragmented and confusing for end users (see Section 2 and 3).

A **set of principles** can help guide both high-level decision- and policy-making and implementation and planning at a more granular level. These principles can cover planning and strategic decisions, the quality of data outputs, management of data assets, and ways of working (such as clear communication and taking a consultative and participative approach).

**Upholding data quality standards** can reduce the burden on end users, improve uptake of data and increase confidence in data. Data quality can be approached from the perspective of:

- **Broad principles and frameworks**, to set out the criteria for managing data well, narrowing down options in areas such as technologies used, licensing, metadata, archival. Principles tend to work in multiple contexts, so a context-specific 'profile' of a set of principles can be helpful for guiding implementation and measuring progress. Examples include: the [UK Government Data Quality Framework](#), the [FAIR Principles](#)<sup>39</sup>; the [W3C Data on the Web Best Practices](#).
- **Macro measures**, for assessing data quality at a national or organisational level, across multiple axes, including accessibility, coverage, completeness. Examples include: the [European Data Maturity Report](#); the [Global Data Barometer](#); surveys

---

<sup>37</sup> Rininta Putri Nugroho Anneke Zuiderwijk Marijn Janssen Martin de Jong (2015), "A comparison of national open data policies: lessons learned", Transforming Government: People, Process and Policy, Vol. 9 Iss 3 pp. 286 - 308

<sup>38</sup> Open Government Partnership, [Open Data to Improve Transparency and Growth](#)

<sup>39</sup> Schultes, Erik, et al. "Reusable FAIR implementation profiles as accelerators of FAIR convergence." International Conference on Conceptual Modeling. Cham: Springer International Publishing, (2020).

of data published in a particular region such as that of [available open data in Scotland](#) produced by Code the City.

- **Micro measures**, which can be a powerful tool for assessing the quality of individual datasets, ranging from generic measures of technical quality to very targeted and domain-specific measures informed by specific user needs. Examples include: the [5\\* Open Data](#) framework<sup>40</sup>; bespoke tools for particular data standards<sup>41</sup>.

**Using technical standards** can reduce the burden on data consumers, and make data publication easier (particularly if relevant support and resources are available).

Technical standards can cover:

- **Using common formats and access methods**, which make open data more accessible and useful to end users, who can integrate data into existing workflows and use common tools and software libraries. Examples include: [CSV](#), [JSON](#) and [JSON Schema](#), the [UK Government list of approved technical formats](#) for the exchange of information; the FAIR Principles suggest the use of common and open formats and protocols for publishing and exchanging information; the [Open Geospatial Consortium library of standards](#) for interoperable geographic information; [Open API](#); the [Realtime Paged Data Exchange](#) (RPDE) format.
- **Metadata (data about data)**, which is essential for providing context for published data and improving both automated and manual discoverability. Metadata can be used to provide information on the content, format, themes, status, ownership and provenance of datasets, and can greatly improve the trustworthiness and usability of published data. Examples include: [DCAT \(Data Catalog Vocabulary\)](#) (Data Catalog Vocabulary); [EU Metadata Quality Assessment](#).
- **Data standards**, to provide an agreed format and requirements for publishing data on specific topics. Data standards enable interoperability and exchange between systems, promote comparability between datasets, and encourage publication of data by making it easier to prepare and share data. Data standards

---

<sup>40</sup> This is potentially problematic due to its very opinionated nature and lack of room to acknowledge real improvements to 'average' data. For example, the transition between two- and three-star data involves shifting tabular data from proprietary to non-proprietary formats but does not account for the enormous potential variance in data quality and utility within a tabular dataset.

<sup>41</sup> Open Contracting data can be assessed using an [online tool](#) and against [key criteria](#) developed against specific use cases. The data quality of major publishers to the International Aid Transparency Initiative (IATI) Standard is compared in a [regular index](#) organised by Publish What You Fund, generated via an open methodology.

are particularly useful for describing data in complex fields or where collaboration and coordination are desired outcomes.<sup>42</sup> Examples include: [360Giving](#) (for grants), [Open Referral](#) (for human services), [International Aid Transparency Initiative](#) (for humanitarian and development assistance) [Open Contracting](#) (for procurement) and [Infrastructure Transparency Initiative](#) (for infrastructure).

- **Using common components and formats** to aid interoperability and reuse, particularly for high value fields like organisational identifiers or very common fields like dates (e.g. [ISO 8601](#)) and monetary values. Using shared concepts and definitions reduces the burden on publishers (who don't have to invent a methodology or definition from scratch) and on users (who don't have to disentangle the meaning of a field on a per dataset basis). Examples include: [org-id](#); [Legal Entity Identifier \(LEI\)](#); currency and country codes; geographic information and coordinate systems (for example, WGS84 coordinates); the [Statistical Data and Metadata Exchange](#) (SDMX) model.

There are also a variety of approaches to specifying technical standards, summarised below.

**Discoverability and availability** of data is critical if it is to be reliably found and used, both manually by humans and programmatically by machines. The most important approaches to making this happen include: the use of accurate metadata and other markup to describe datasets and improve search-based discoverability; clear policies around uptime and retention; and the use of centralised services to collate publications and provide context, signposting and quality control.<sup>43</sup>

The Scottish Government uses the FAIR Principles, which are very useful but primarily concerned with technical quality and design. This leaves gaps around how to prioritise and plan data publication, how often data will be updated and how long it will be available for, and how publishers should interact with users. It is vital to ensure existing and future policies around data publication have buy-in from internal and external stakeholders, and to encourage uptake by integrating the practical aspects of these principles, along with expectations around quality and measuring progress, into organisational data strategies.

Discoverability and availability of data in Scotland is an area of weakness due to fragmentation, uneven publication policies among public sector bodies, a lack of

---

<sup>42</sup> Open Data Services, [Standards Lab Handbook](#)

<sup>43</sup> Davies, T, [Focussed Futures: the Data Portal as...](#), (2023)

signposting, and lack of reliable access and retention to data.<sup>44</sup> This is reflected in, for example, the fact that civil society created a grassroots data portal in the absence of an official alternative and that awareness of successful projects to publish complex data sets (for example, around procurement or fiscal data) remains low, even among open data enthusiasts. Discoverability is now being partially addressed via a CivTech challenge to provide accessible search and quality assessment of open data in Scotland.<sup>45</sup>

### 2.5.3 Tools, resources and support

Practical guidance in the form of tools, resources and support can help publishers identify and achieve key activities and help users engage with and generate impact from published outputs.

#### Information and documentation

Publishers and users both benefit from clear guidance on what open data is and the key processes and considerations involved in producing and using open data. These types of resources may include non-technical explainers, technical documentation and practical tools or templates to assist with common activities in the data value chain. This can take the form of:

- Briefings on the value of and requirements for impactful open data for non-technical audiences.<sup>46</sup>
- General open data toolkits that combine documentation and signposting to other resources.<sup>47</sup>
- Guidance on specific open data topics, such as licensing, anonymisation or identifiers.<sup>48</sup>
- Publisher-facing documentation of tools, operational processes and activities that are part of the data value chain.<sup>49</sup>

---

<sup>44</sup> Watt, I. [What Is Open Data and Why Does it Matter?](#), David Hume Institute, (2023)

<sup>45</sup> Krishnadas, G. [The evolution of our data discoverability solution Dtechtive through the CivTech programme of the Scottish Government](#), 2023

<sup>46</sup> Lord, J. and Kiepe, T. Structured and interoperable beneficial ownership data, Open Ownership, (2022)

<sup>47</sup> A number of countries have created open data guides and toolkits in this vein, aimed at addressing common challenges faced by publishers. Kucera, Jan, et al. "Methodologies and Best Practices for Open Data Publication." DATESO. 2015. The Scottish Government has its own open data resource pack that, while somewhat out-of-date, does tackle some of the major issues.

<sup>48</sup> Dodds, L. Open Data Institute, Publishers Guide to Open Data Licensing, (2013)

<sup>49</sup> Interagency Standing Committee, Operational Guidance on Data Responsibility in Humanitarian Action, 2023.

- User-facing documentation of datasets, access and analysis tools, and underlying methodologies.<sup>50</sup>
- A knowledge base that collects and documents common queries and provides guidance or context for using and interpreting open data.<sup>51</sup>
- Collating and signposting existing resources, which may otherwise be fragmented and difficult for users to discover, and by designing user journeys through such resources.

### **Tooling and infrastructure**

Some central provision of technical resources, tooling and infrastructure to aid open data publication can provide significant savings, quality control and generally ease the road to publication. This can take the form of:

- Provision of a central data portal and/or publishing service to avoid fragmentation and divergence in practice.<sup>52</sup>
- Designing dedicated systems for analysing, visualising and explaining data on particular themes.<sup>53</sup>
- Resourcing software or software-as-a-service tools to make data publication or analysis easier.
- Providing data quality tools to identify and fix issues with the data or metadata.
- Using open source methodologies and licensing to promote the reuse of tools and techniques and give users visibility on changes.<sup>54</sup>

### **Direct, hands-on support**

Expert support on publication can be very helpful, especially when responsibility for publication is distributed and organisations may lack the specialist skills or resources needed to meet their obligations. Publishing novel or complex datasets, or publishing to data standards, is another situation where direct support can be useful. This could involve:

- A helpdesk model where publishers and/or users can resolve queries and receive direct technical assistance.<sup>55</sup>

---

<sup>50</sup> The UK's Office for National Statistics publishes an extensive collection of methodology documents for its regular publications

<sup>51</sup> The World Bank maintains an extensive knowledge base about its open data products and methodologies.

<sup>52</sup> Many countries offer a data portal, which can serve multiple roles in the data ecosystem. Research by the Open Data Institute collated a number of design patterns, and anti-patterns, for data portals.

<sup>53</sup> The Scottish Index of Multiple Deprivation has a dedicated mapping tool.

<sup>54</sup> Companies House, for example, develops its APIs for open company information on Github.

<sup>55</sup> The helpdesk model is used by most large multi stakeholder open data initiatives to facilitate publication in line with data standards or other quality requirements. Open Data Services, The Helpdesk



- Procurement of external resources to execute complex or novel data publication initiatives.
- Funding permanent public sector staff either directly in publishing organisations or to assist publication across a number of public sector organisations.
- Secondment of staff to organisations that lack experience or capacity to publish open data.

### **Community and capacity building**

Open data publication works best as a collaborative endeavour. Guidance should reflect this by cultivating a supportive community that can reflect honestly on the challenges and find collective solutions. This can be achieved by:

- Developing standards and tools in open source repositories, where users can provide feedback and report issues.
- Creating and facilitating forums, or other online spaces, where publishers and users can share experiences, ask questions or provide feedback.<sup>56</sup>
- Hosting facilitated community calls or office hours where substantive issues can be discussed.<sup>57</sup>
- Offering training sessions and self-service training materials to improve understanding of specific issues or build skills.

Scotland already has many of the components of a useful support ecosystem in place. The challenge will be to iterate on these, while pulling them into a coherent and comprehensible package for data publishers and end users.

---

Model. The World Bank runs a user-facing helpdesk to promote the use and reuse of its open data publications.

<sup>56</sup> IATI and Open Referral have healthy online communities of practice that provide ongoing dialogue between various stakeholders. Companies House maintains a more technically-oriented forum for users of its company data products.

<sup>57</sup> 360Giving and OpenActive both run regular onboarding and adoption calls.

## 3. Open data in Scotland

### 3.1 Policy environment

Scotland has a number of strategy initiatives directly or indirectly related to open data:

- [Freedom of Information \(Scotland\) Act](#), 2002: An Act of Scottish Parliament which entitles any person to be given information held by a Scottish public authority upon request, subject to certain conditions and exemptions.
- [INSPIRE \(Scotland\) Regulations](#), 2009: An implementation of EU Directive 2007/2/EC which requires Scottish public authorities to make spatial datasets (for example, map data) available.
- **The UK and Scottish [Statistical System](#)**: Official statistics in Scotland provide an accurate and up-to-date picture of the economy and society. There is a [Code of Practice for Statistics](#) that sets the standards that producers of official statistics should commit to. Compliance with the Code gives users confidence that published government statistics have public value, are high quality, and are produced by people and organisations that are trustworthy. Publishing data openly is an important component of the Code of Practice for Statistics.
- [Open Data Strategy](#), 2015: Scotland's Open Data Strategy established the principle of 'open by default' and aimed to promote open data in order to improve the delivery of public services, catalyse social economic benefits through innovation and reuse, and improve transparency and accountability of the public sector. The strategy is supported by guidance and a toolbox.
- [Scotland's Open Government Partnership \(OGP\) Action Plan](#), 2021-2025: Scotland is a local member of OGP. The OGP is a global partnership between civil society and government committed to transparent, participatory, inclusive and accountable governance. Data and digital is one of five commitments made as part of Scotland's OGP action plan. It includes plans to open up datasets relevant to other commitments, and to develop the open data infrastructure.
- [A Changing Nation: how Scotland will thrive in a digital world](#), 2021: Scotland's Digital Strategy underlines Scotland's commitment to thriving in a digital world. The strategy identifies a number of priority areas where digital

initiatives can contribute: improving the delivery of public services, economic recovery and growth, addressing climate change, and improving digital skills and connectivity to ensure participation. The strategy identifies open data as a key implementation tool and commits to making more open data available and discoverable.

- [Scotland's AI Strategy: Trustworthy, Ethical and Inclusive](#), 2021: Scotland's AI strategy sets out the goal to make Scotland a world leader in trustworthy, ethical and inclusive AI. The ongoing supply of high-quality and responsibly-produced data is identified as one of the foundations for success, and open data as a key tool to make this happen.
- [Greater access, better insight, improved outcomes: a strategy for data-driven care in the digital age](#), 2023: Scotland's first digital strategy for health and social care is aimed at improving how data is used in the sector to enable the public to access and use data to improve their health and improving how health and social care outcomes and services are delivered. The strategy identifies the need for useful, relevant, and accessible data to be made available openly, within a strong framework that ensures data is shared appropriately.
- [Participation Framework](#), 2023: Scotland's Participation Framework provides guidance and tools to improve public involvement in the design and delivery of policy and services. The framework was delivered as part of Scotland's OGP commitments. The framework identifies the publication of data as a tool to enable informed participation. The tools and techniques in the framework are also relevant for planning open data publications that meet user needs, build trust and ultimately generate impact.

In addition, there are a number of practical initiatives directly or indirectly concerned with improving open data publication, use and impact:

- The Digital Directorate hosts the '**Better Data Community (Making Data Better, Smarter and More Open)**'. This community of practice is intended for anyone in the public sector who is interested in improving the quality of the data collected in the public sector, how it can be more effectively managed and reused. It allows practitioners to share best practice for the publication of open data.
- The [Digital Data Maturity program](#) has created a **Data Transformation Framework** that will increase the digital capability of public sector organisations.

The program has run several cohorts and is now trialling a self-service approach.

- Outside the public sector, [Code the City](#), a charity that campaigns for better open data, hosts a Slack community, an annual conference, and its own data portal.

## 3.2 Publication landscape

Open data-related roles and responsibilities provide important context for assessing the state of open data publication and use. In Scotland, there is a key distinction between:

- **Policy** responsibility, which sits with the Open Data Team<sup>58</sup> in the Data Division of the Digital Directorate of the Scottish Government; and,
- **Publication** responsibility, which is delegated to potentially hundreds of public sector bodies, including local authorities, government departments and organisations, the Scottish Parliament, and NHS Scotland.

This highly dispersed responsibility for open data publication in Scotland is not unusual but it does present a dual challenge to:

- Translate the enthusiasm and direction of travel from the Scottish Government into publication by public sector organisations; and,
- Guide current publication efforts into a set of coordinated publication practices that will result in a coherent, accessible and impactful data publication landscape across Scotland.

These challenges are analogous to those faced by large multi stakeholder data initiatives that rely on the distributed publication of data that would otherwise be locked up in organisational silos to generate a useful corpus of shared data. Scotland may benefit from adopting a similar high-level approach to coordinating distributed publication at scale to that used by these initiatives, which combines advocacy and strategic work, ongoing outreach and partnership-building, and the provision of practical resources and support. This would imply some intentional blurring of the line between policy responsibility and publication responsibility.

---

<sup>58</sup> This team also manages [statistics.gov.scot](https://statistics.gov.scot).

Scotland's results on open data to date have been mixed. Scotland has some excellent open data, including geospatial data, and key resources like Scotland's Census and Official Statistics that are both trusted, accessible and well-used. In addition, huge progress has been made on the publication of data in complex domains like fiscal transparency. But that said, the coverage of open data is uneven (and particularly low among local authorities) and there is a sense that progress and enthusiasm has stalled.<sup>59</sup>

### **3.3 Findings from workshops**

The desk research was complemented with a series of public workshops to discover attitudes towards open data among technical and non-technical practitioners and stakeholders in Scotland. The workshops, run by Fractals Coop, were designed to provide an accessible and open forum for a wide range of participants to talk about their experience of, and hopes for, open data. Participants in the workshops were drawn from a range of sectors including central and local government, other public sector, civil society, academia and the private sector, with a variety of job roles and technical or policy expertise represented. Unsurprisingly all attendees had some experience with open data, and were therefore not fully representative of the much larger pool of those who might benefit from open data in the future.

---

<sup>59</sup> Watt, I. [What Is Open Data and Why Does it Matter?](#), David Hume Institute, (2023)

**Figure 3. What does open data mean to you? Findings from the workshops grouped by theme**



Source: Scottish Government, Summarised findings of workshops run by Fractals Coop, in conjunction with Open Data Services

Several important themes emerged from these workshops:

- **Policy:** Participants emphasised the need for clearer direction on open data and on adopting open data policies that are clear, implementable and enforceable. Participants noted the drawbacks of the unenforced 'open by default' approach and suggested combining incentives for good practices, support for implementation, and accountability for those not meeting their obligations. Participants also noted that open data could be attached to broader policy initiatives and strategies, including on AI and digital, and that alignment with international practices, such as the EU Open Data Directive, would be helpful.
- **Mindset and Culture:** Participants identified cultural and mindset barriers in the public sector as a significant barrier to open data publication. Publishing open data is perceived as risky with few tangible benefits for producers and, as a result, there was a noted reluctance to share data openly and a need for a cultural shift towards more collaborative and open practices. Concerns about data publication errors, including incorrectly publishing personal or commercial data, and the need for clear guidance and training were also emphasised. It was

noted that data can be an off-putting and overly technical term, which narrows the conversation on openness.

- **Needs of Data Producers:** Participants noted that data producers face numerous practical challenges when publishing open data, such as skills and resourcing gaps, a lack of practical support, and gaps in enabling infrastructure such as a centralised repository for open data. Participants noted the helpfulness of existing guidance and the need for updates and expansion to such material, to include standardised tools and practices for data sharing and interoperability and to cover key topics such as data anonymisation techniques.
- **Needs of Data Users:** Participants identified several key areas where data users encounter difficulties with open data. These include: discovering and accessing relevant information, finding data that is high quality and up-to-date, understanding the context and potential applications of the data, and having the right skills to access and use data. There was also a call for more user-friendly systems and tools that can facilitate easier interaction with and analysis of open data, especially for non-technical groups.
- **Impact and Value:** Participants were generally enthusiastic about the potential for open data and stressed the need to shift wider perceptions by demonstrating the economic and social benefits. Participants suggested that showcasing successful use cases and real-life applications could inspire greater adoption and innovation and noted the need for mechanisms to measure and communicate the impact of open data initiatives to justify ongoing investment.

These insights highlight many of the important links in the open data value chain model (Commit-Build-Use-Learn) discussed in section 2.4 and the issues that stop open data in Scotland from reaching its full potential.

### 3.4 Opportunities for the future

Scotland has an opportunity to take advantage of a policy environment which is amenable to open data publication, and enthusiasm and capacity in the community and private sector about the potential of open data, as a strong foundation for driving forward the next phase of impactful open data implementation.

The key will be to identify areas of strength and capacity within the Scottish Government, and define carefully the role that the Scottish Government will play in the wider landscape. It is important to set expectations about responsibilities and outcomes,

and to prioritise an approach that will allow the Scottish Government to have the most impact, making best use of limited resources available.

This is a direction setting and coordination role, with responsibility for the actual implementation distributed among other stakeholders. The Scottish Government can, for example, advocate internally to ensure the culture and expertise is conducive to meeting high level policy goals; identify specific areas where open data would have a significant impact; convene communities of practice and help to coordinate and intentionally direct their activities; identify and commission appropriate support for data publishers; monitor, measure and celebrate the benefits created by open data projects. This has the potential to create a positive feedback loop, where demonstrating impact increases enthusiasm and buy-in internally and externally, setting the stage for further improvements and innovations in the long-term.

The following section highlights some more specific findings from the workshops, linking these to detailed recommendations on improving the impact of open government data in Scotland.



## 4. Recommendations

### 4.1 Set out a vision that makes the case for open data

#### What the research found

- Widespread excitement about the potential for open data in Scotland.
- Weariness, wariness and uncertainty about how to win the argument for open data and to turn commitments into reality.
- A desire to see open data publication efforts translate into real impact for a diverse set of stakeholders, held back by the marginalisation of open data as a technical project.
- A potential conflict between recent research suggesting that targeted effort is the optimal route to impactful open data and the ‘open by default’ principle underlying Scotland’s existing open data strategy.

Our first recommendation is to **establish a clear and compelling vision for open data** that positions open data as enabling infrastructure that supports people and organisations across Scotland to access high-quality information. This vision should be aligned with wider digital and policy priorities and explicitly connect the publication of open data to mechanisms of change and improvement. The vision should be shared, understood and promoted by senior stakeholders to counteract the perception that open data is a ‘nice to have’ rather than a priority or obligation and to give implementers the confidence to publish.

Much of the groundwork for this is already in place, and the [consultation on data standards](#) in Scotland’s public sector identified a very similar need. Any vision should align easily with the principles and priorities outlined in the 2021 digital strategy, [A Changing Nation](#), which already sets out how open data initiatives can complement and contribute to broader digital objectives. Both the digital strategy and [Scotland’s National Performance Framework](#) point towards an alignment with broader policy objectives. Scotland’s Open Government Partnership commitments situate data as an enabler for technical and non-technical audiences who need high-quality information to advance policy objectives across a number of themes.

The following sub-recommendations are designed around producing a compelling and achievable vision.

### 4.1.1 Broaden the discussion beyond technical outcomes

The research showed an underlying frustration with an apparent focus on technical outputs (i.e. the existence of a particular dataset in a particular format) to evaluate the success of open data publishing, rather than the outcomes that stem from the availability of high-quality information to anyone who needs it. Thinking and talking about information-availability rather than data may be helpful to:

- Reduce the risk of open data being pigeon-holed as a technical project serving a niche technical audience.
- Highlight the gaps between data availability and information access and use.
- Identify user needs around information access and translate these into technical activities in open data programs.

Widely-available, high-quality information will of course continue to rely on solid technical implementation, as later recommendations discuss in detail, but broadening the discussion on purpose will make it easier to justify technical work.

### 4.1.2 Take a purpose-directed approach to open data publication

Our research suggested that the idea of open data and the principle of open by default are both widely supported. But these abstract and all-encompassing concepts can be overwhelming for publishers (who need to decide what to focus on and in what order) and for open data advocates (who are less likely to make focused arguments to open up specific datasets for specific purposes).

We would therefore recommend that this vision for open data in Scotland reflects the more targeted and attainable approach to publication seen in the so-called [third wave of open data](#), which ‘seeks not simply to open data, but to do so in a way that focuses on impactful reuse, especially through inter-sectoral collaborations and partnerships’.<sup>60</sup> This should not be seen as a rolling back of ambition from ‘open by default’ but, instead, as a strategy to maximise impact in an environment where public sector publishers face capability and capacity constraints and where data users may benefit more from targeted and concerted effort in specific areas.

---

<sup>60</sup> Scotland’s OGP commitment on financial transparency as a good example.

### 4.1.3 Agree thematic priorities and be explicit about how impact is expected to happen

A purpose-directed approach to data publication will require choices about where to focus effort, based on an understanding of how these efforts are expected to deliver impact. We therefore recommend that a vision for open data includes some high-level priorities for open data (while leaving some space for emergent needs) and is explicit about the mechanisms that will connect publication to impact. Together these should allow more effective consultation, planning and implementation.

The digital strategy, the 2015 [open data strategy](#) and our literature review identify a number of **mechanisms** via which open data can have a positive impact. These can be summarised as: contributing to growth and innovation; delivering social value; building trust in government and public services; improving public sector delivery and coordination; and, enabling collaboration between government and other stakeholders. Scotland's [Open Government Action Plan](#), meanwhile, includes a strong **thematic** focus on specific data areas, including financial transparency and the environment. Explicitly combining these approaches to document the expected impact of open data publication may be helpful in defining priorities, discovering neglected or high potential areas, and taking a purpose-directed approach to data publication.

### 4.1.4 Collaborate and communicate on the development of priorities

Our research uncovered a certain level of distrust and cynicism around the Scottish Government's attitudes to open data. There is a risk, therefore, that a more focused vision for open data is misinterpreted as a turn away from openness, and leads to disengagement and fatalism among stakeholders. Building a coalition around a purpose-directed approach will be one of the key challenges, and requires building trust and dialogue with those outside the public sector. That dialogue will also be crucial to the success of purpose-directed publication, which requires a deeper understanding of user needs and priorities in order to direct resources where they will be most impactful.

### 4.1.5 Make space for high-ambition goals and initiatives

Our research uncovered a degree of frustration that potentially significant pieces of open data infrastructure may not be available or accessible (e.g. a land information system or an openly-licensed address dataset).<sup>61</sup> Such potential big ticket open data

---

<sup>61</sup> Postcode to area lookup files are made available in the [Scottish Postcode Directory](#). However property references and address lookups require access to proprietary databases. Owen Boswarva, [Thoughts on](#)

projects are particularly complex because they rely on significant cross-departmental collaboration, a change in operating model from a data holder, regulatory changes, or digitalisation. They therefore tend to fall down the agenda in short term planning cycles. It is crucial to keep big ideas on the radar, however, as these may become crucial open data assets in the future - and an inspiring and trusted vision for open data in Scotland may create favourable conditions to realise such projects.

## 4.2 Embrace long-term collaboration and commitment

### What the research found

- A disconnect between the high expectations for open data publication and the government's ability to increase the quantity and quality of data, while making publication simple and routine.
- The effort involved in publishing data sustainably and at high quality is rarely visible, particularly in terms of new and complex datasets.
- Publishers are often disconnected from users and potential users of data, creating a producer-consumer dynamic at odds with the need for collaborative approaches to maximise the impact of open data.

Our research found a disconnect between the complex reality of how open data gets published and the expectation that the government has simple levers that could broaden publication and make publication simple, routine, and high quality - and do so quickly.<sup>62</sup> Setting expectations too high, and failing to celebrate success, creates a risk of disillusionment among publishers and of disempowering those stakeholders outside government whose input and cooperation is needed to make open data genuinely impactful.

It may therefore be helpful to **reframe open data as a collective and collaborative commitment to a long term project** in order to change the way the public sector thinks about open data and the way that those outside government think about open data production by the public sector. A more realistic narrative on open data publication is likely to contain several elements:

---

[campaigning for open address data in the UK](#), (2023). Andy Wightman, [Scotland's Land Information System – what is it and why does it matter?](#), David Hume Institute, (2023).

<sup>62</sup> These levers do exist and can drive genuine improvement, as discussed in Recommendation 3.

- Impactful open data publication should be seen as achievable at reasonable cost, with low risks and significant benefits for publishers and external stakeholders.
- The complexity of open data publication, including the policy, governance and technical work involved, should be visible both within and outside government.
- Open data publication should be viewed as an ongoing collaborative process between data publishers and users, rather than a one-off event, and as a way to iterate towards impact.
- Rather than an isolated project, open data is a component of several broader trends in government, including towards openness, participation, and digital maturity - and, while open data can act as a catalyst, it is likely to gain the most traction where these trends are strongest.

Recommendation 3 sets out how this narrative can be made concrete and credible through a realistic and transparent implementation strategy.

### **4.3 Create an implementation strategy that is intentional, realistic, transparently monitored, and celebrates success**

#### **What the research found**

- Uncertainty about what the Scottish Government was doing about open data and what it expected others to do.
- An uneven pattern of progress based on organisational factors.
- Negativity about the state of open data in Scotland, particularly from civil society, and a downplaying of progress and achievements in open data.
- The difficulty of finding consistent documentation of plans and progress relating to open data.
- Low awareness of success stories in open data and access to information.

Our research has unveiled several challenges in translating commitments on open data into reality, combined with a low awareness of the significant progress that has been made even among very engaged stakeholder groups. Experience in other countries suggests that the development of concrete and measurable plans has been helpful in

motivating and monitoring data publication. In addition, any shift to purpose-directed publication and long-term commitments to open data needs to be balanced with mechanisms to keep up momentum on publication and data quality, and to unlock impact and value as soon as possible.

We therefore recommend the creation of an **implementation strategy for open data in Scotland that is realistic, transparently monitored, regularly updated and designed to celebrate success**. The following sub-recommendations are in support of such a strategy.

### **4.3.1 Targeted and intentional data production**

Targeted production of high value datasets or data that serves current priorities, over the potentially scattergun 'open by default' approach. This should ensure that open data releases serve concrete purposes and meet defined needs and that resourcing can be directed where it will do the most good. Any data that cannot be made fully open should be justified and transparently communicated.

A national strategy can identify high level priorities and particularly complex data publication projects but there should also be guidance for individual publishing organisations who need to plan their own publishing activities.

### **4.3.2 Sustainable open data publication**

Adopting the principle of sustainable open data publication, with clear expectations set around the lifecycle of the data, including metadata about publication history (a changelog) and transparent communication and a roadmap about the future of the content or availability of datasets. Sustainability also requires ongoing demand, which requires that the data meets quality standards related to coverage, timeliness, accuracy, and interoperability (as discussed in Recommendation 4). The parallel development of publishers' data maturity will be integral to this effort.

Take a long-term and iterative approach to open data work that emphasises ongoing improvement over one-off data publication. This approach allows for value to be realised early, while ensuring that data quality and other factors improve over time. A long-term approach will also require goals and targets that recognise the importance of improvement in data (discussed further in Recommendation 4).

### **4.3.3 Review, report, celebrate**

Build momentum, trust and accountability through regular reviews and transparent reporting. These plans should be publicly available and easily discoverable. Transparency in progress and accomplishments should extend beyond official OGP plans, creating a comprehensive view of open data initiatives' impact and advancements in Scotland.

Apply standard assessments of the benefits and risks of publishing open data and standard techniques to reduce the risk of disclosing personal or confidential data. Risk aversion was highlighted as a key organisational barrier to publishing open data and risk reduction as a time-consuming and stressful process. Adopting standard methods for assessing and managing risk (for example, techniques for anonymisation, criteria for exclusion or redaction, and clear redress mechanisms) could help here. Making such assessments public could also help build trust in data. The [data nutrition label](#) is a useful example of how this could work.

Ensure that success is highlighted and celebrated, both in formal reporting and through other fora that connect publishers, users and other practitioners. Some success is likely to be measurable but it is equally important to identify and tell the human stories behind data publication and impact.

Recommendation 4 sets out a more granular view of how to support the technical activities needed to implement such a strategy.

## **4.4. Set the standards for open data publication**

### **What the research found**

- Publishers appreciate the existing technical guidance resources but are still left with many large and small implementation decisions.
- Less choice and more direction would help publishers and give data users more confidence.
- Users struggle to discover and understand the scope of open data.
- Use and reuse of data is held back by uncertainty over licensing, sustainability, and data quality - and a lack of specialist resources holds back improvement in these areas.

Our research uncovered significant challenges for open data publishers in Scotland, from the technical challenge, the often overwhelming number of choices publishers are faced with, the perceived risks of publication and uneven data governance processes. These challenges reduce the number of datasets that are published and reduce the utility for users.

We recommend that the government provides clear direction, specifying quality standards for open data publication that incorporate technical best practice, usability and findability, and clear data governance, in line with the [UK Government Data Quality Principles](#).

The following sub-recommendations outline how to provide publishers with clearer direction and establish quality standards for open data publication, bolstering the confidence of data users and advancing the effectiveness and impact of open data in Scotland.

#### **4.4.1 Tailor data quality principles to the context**

The [FAIR Principles](#) (Findable, Accessible, Interoperable, and Reusable) are already integral to thinking on Scotland's public sector data and the detailed guidance notes provide a useful starting point for building assessments of and targeted support for open data publication. As the needs around open data become clearer, and there is increased justification for practice to converge, it may be useful to consider a [FAIR implementation profile](#) (FIP) for the public sector.

Adopt new measures of data quality and coverage in place of or in addition to the target of 5-star open data. The [5-star open data system](#) is useful when information is not published at all, but is an increasingly blunt instrument when trying to improve data quality to unlock specific use cases. There are several alternatives that, depending on context, may be more useful. These include macro-measures of data availability, quality and usability in particular jurisdiction, as used in the EU's [Open Data Maturity Report](#); functional assessments of data quality and availability on specific themes, as used in the [Global Data Barometer](#); and, domain-specific or standard-specific assessments of individual datasets, as used in the [Aid Transparency Index](#).

#### **4.4.2 Mandate the use of technical standards.**

As part of a drive to reduce the burden of publication and improve interoperability across datasets, it may be useful to adopt a more proactive stance on standards for publication and exchange formats, common data components, specialist data standards



for complex data, and metadata standards and vocabularies. Adherence to such standards should be measured in any assessment of data quality and publishers should be able to easily test their implementations by validating against the standards. There is a detailed review of this area in [Section 2.5.2 of this report](#).

#### **4.4.3 Make discoverability an intrinsic part of publishing data**

Users should be able to find data that they know about and discover data that may be of interest to them - and the experience should be relatively consistent across publishers. The difficulties reported by users in finding relevant data are reflected in the creation of an informal open data portal by Scotland's open data community and in a government project to improve findability via a CivTech Challenge. Resolving this issue, and ensuring that users can take multiple routes to find a dataset, could solve a significant pain point. Potential approaches include: ensuring that metadata is present; adding markup to publication pages for search engine discoverability (for example, conforming to [Google's dataset structured data requirements](#)); requiring visibility at a central location via direct publication or by aggregation; and, ensuring that users can both browse and programmatically interact with the catalogue of published datasets.

Recommendation 5 sets out how to provide ongoing support to new and established data publishers and users.

### **4.5 Provide targeted technical guidance and support for publishers and users**

#### **What the research found**

- Existing technical guidance resources are useful but could be expanded to cover more of the publication process and provide clear guidance at specific decision points.
- Data publishers can be isolated during planning and implementation.
- Users can struggle with the purpose, scope and subtleties of open datasets.

We recommend providing targeted technical guidance and support for data publishers and data users, and doing so across multiple modalities. The following sub-recommendations aim to entrench the idea that open data is a long-term, collaborative

project, improve confidence in the data itself, and, over time, improve trust in the government's strategic support of open data.

#### **4.5.1 Self-service support for publishers**

Expand and update the self-service toolkit for publishers. The open data resource pack within Scotland's Open Data Strategy includes valuable reference material and guidance, and an update would be helpful to bring this in line with current priorities and the learning since publication. In particular the practical tools listed in [Annex A of Scotland's Open Data Strategy](#) could be expanded to address current pain points. This might include:

- An information mapping tool to assist with planning complex data releases. This could help identify, for example, the components of a planned dataset, any required cleaning or merging, current ownership of the data and data governance policies, any technical details of how the data is stored and accessed, and stakeholders involved.
- A standards catalogue to help identify existing formats for publishing data.
- A list of tools and libraries to assist with publication, including automation of data production and testing.
- Templates for data risk assessments, data sharing agreements and other data governance processes.

#### **4.5.2 Direct support for publishers**

Many publishers are starting from scratch with open data publication, in terms of institutional knowledge and capability. Some support provision could therefore be worthwhile to reduce the burden on individual publishers and to avoid reinventing the wheel and divergent practice. The appropriate type of support is context-dependent but could include:

- A helpdesk that answers queries and provides asynchronous support to data publishers. This is a common service model for large data initiatives where responsibility for publication is highly distributed.
- Drop-in sessions for publishers.
- Live and on-demand training on specific issues in data publication.
- Provision of technical infrastructure to assist with self-service publication.

### **4.5.3 Join up communities of practice**

Join up communities of practice and bring publishers and users together. Scotland has several communities of practice in open data and adjacent areas, including Open Data Scotland, the OGP community, organisations in or graduated from the Data Maturity Program, and the government's own data community of practice. These communities are valuable assets that could be leveraged more. For example, peer support and validation can be crucial for removing the fear factor from open data publication and helping to spread best practice, so providing more informal channels to discuss open data may be helpful. And there is currently no mechanism to bring producers and users of data together to design improvements, discuss priorities and uncover potential demand; establishing a multistakeholder data user group to address this.

### **4.5.4 Lightweight support for data users**

Recommendations related to technical standards in Recommendation 4 inherently improve data comprehensibility and reusability when implemented well. Recommendations related to reporting and celebrating in Recommendation 3 will increase awareness of the availability and potential of open data, so additional support for data users at this stage can be lightweight.

Support for data users can take the form of context, explanation and points of contact for published data. Having data that is comprehensible to non-technical audiences is a strong marker of maturity in open data publication. Data can be difficult to understand if it is just a spreadsheet or other technical artefact. The barrier to entry can be reduced through good documentation, non-technical explanations of the dataset (e.g. [ONS bulletins](#) that accompany statistical releases), and demonstrations of data use.

## 5. Summary and next steps

The Scottish Government can establish a position as a central coordinator and set the direction for unlocking the potential of open data in Scotland. Collaboration and long-lasting partnerships with the community and other organisations are critical to making this achievable.

The recommendations on how to do this are as follows:

**1. Set out a vision that makes the case for open data.**

This should centre social and economic impacts over technical artefacts; be purpose-directed with thematic priorities; be clearly communicated to build trust; and set expectations for the long-term.

**2. Embrace long-term collaboration and commitment.**

Partner with community stakeholders and private sector organisations to make best use of limited government resources; make progress visible; align with broader trends towards government openness.

**3. Create an implementation strategy that is intentional, realistic, transparently monitored and celebrates success.**

Target data publication in line with thematic priorities; set expectations around a sustainable approach; regularly review, measure and report on impacts.

**4. Set the standards for open data publication.**

Provide direction on data quality, technical standards and discoverability of data.

**5. Provide targeted technical guidance and support for publishers and users.**

Establish direct and indirect support for data publishers and users; centre existing communities of practice.

The Scottish Government has an opportunity to take advantage of two decades of established best practice around open data publication, and to lean into existing communities of practice, both locally and globally, to establish a sustainable, long-term strategy for generating positive impact using open data in Scotland.

Challenges will be around building trust in both the vision and the implementation. It will be necessary to improve the organisational culture around data publication, providing training to build internal expertise where necessary. Establishing a culture of continuous, iterative improvement and celebrating small successes with regular, transparent reporting may be a shift from previous ways of working.

For the next steps, we recommend:

- Early identification of prospective partnerships and collaborative opportunities.
- Determine how best to measure impact and assess risk, and put processes in place to do so as early as possible.
- Set out responsibility for making decisions about principles and frameworks for data quality and technical choices that will guide data publication, and plan how to communicate those decisions once made with internal and external stakeholders.
- Identify key areas to target with open data initiatives, based on high level government priorities, and local knowledge about where the most impact will be made.



© Crown copyright 2024

**OGL**

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit [nationalarchives.gov.uk/doc/open-government-licence/version/3](https://nationalarchives.gov.uk/doc/open-government-licence/version/3) or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk).

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

This publication is available at [www.gov.scot](http://www.gov.scot)

Any enquiries regarding this publication should be sent to us at

The Scottish Government  
St Andrew's House  
Edinburgh  
EH1 3DG

ISBN: 978-1-83601-623-6 (web only)

Published by The Scottish Government, August 2024

Produced for The Scottish Government by APS Group Scotland, 21 Tennant Street, Edinburgh EH6 5NA  
PPDAS1487558 (08/24)

W W W . g o v . s c o t