Using open data for policy

Summary of training workshop

29 June 2017

Stats NZ
Tatauranga Aotearoa

The Policy Project
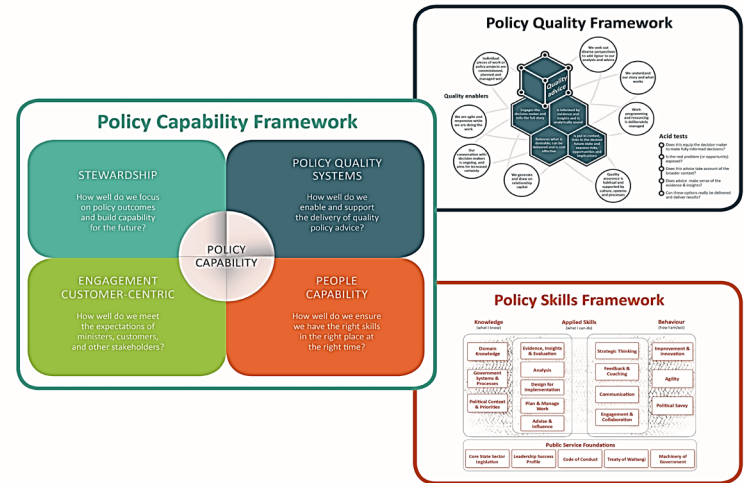Responsive today,
shaping tomorrow

Background and intent

Stats NZ hosted the workshop on how to use open data for policy design. It was facilitated by Ellen Broad (Associate – Policy, Open Data Institute), and supported by the Department of Internal Affairs, who administer the open data platform (data.govt.nz), and the Policy Project.

The Policy Project supports the Head of Policy Profession in working with policy leaders and practitioners to improve policy quality and capability (see Policy Project website for more info). The Policy Project shares a common goal with Stats NZ: to develop capability to provide more evidence-informed policy advice. Mutual aims include to:

- Increase data literacy among policy analysts

- Provide guidance on how policy teams go about using data

- Develop informed customers to drive demand for evidence based policy and ensure they know what to ask for

- Demonstrate the value of data and the policy successes delivered

The intent of this summary is to capture and spread the learnings of the workshop for the benefit of the broader policy community, supporting the above aims.
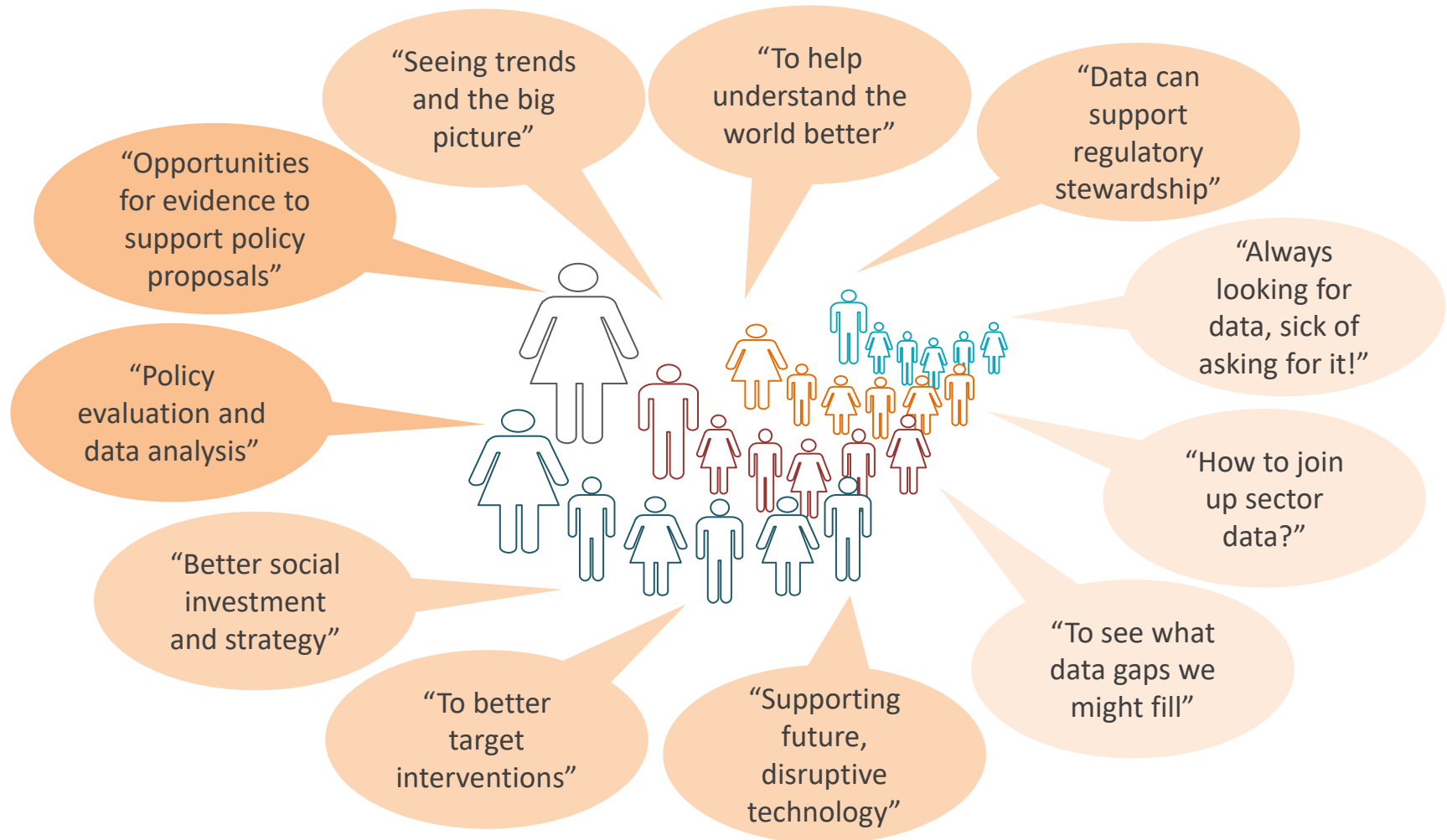
These aims are also supported by the policy improvement frameworks and tools, including:
- the Policy Capability Framework (which emphasises the need for research, analysis and knowledge e.g. good data architecture)
- the Policy Quality Framework (which emphasises the need for advice to be well informed by data, evidence, insights)
- the Policy Skills Framework (which describes related skills e.g. understanding data principles and analysis methodology).

Why does open data matter?

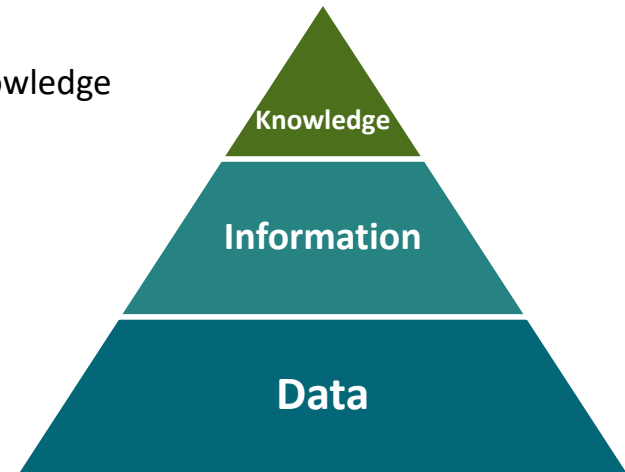When asked what their interest in open data is, participants said…

"Seeing trends and the big picture"

"To help understand the world better"

"Data can support regulatory stewardship"

"Opportunities for evidence to support policy proposals"

"Always looking for data, sick of asking for it!"

"Policy evaluation and data analysis"

"How to join up sector data?"

"Better social investment and strategy"

"To better target interventions"

"Supporting future, disruptive technology"

"To see what data gaps we might fill"

What is open data?

Data is "information without context", including collections of:

➢ **facts**, **information** and statistics that can be analysed to develop new knowledge

➢ **numbers** assigned as values to quantitative variables and/or characters assigned as values to qualitative variables.

> "Whether information is considered to be "data" depends on its use, and its users."

Knowledge

Information

Data

Data can be closed, shared or open. A data set can be publicly available (e.g. a phone book) without being open.

For data to be considered "open", it must:

➢ be **accessible**, which usually means published on the web

➢ be available in **machine-readable** format: predictably structured for programmatic consumption e.g. CSV file

➢ have a **licence** that permits anyone to access, use and share it (commercially and non-commercially).

Open data is data that anyone can access, use or share.

"Open government data" has additional features, including:

➢ **non-personal, unclassified and non-confidential**

➢ collected, commissioned and created by government agency

➢ publicly funded (or with agreement of parties involved – best to communicate intention to make data open upfront)

See here for more information on open government data and New Zealand's Data and Information Management Principles, here for how to keep data safe e.g. remove personal identifying information ("de-identification"), and here to determine the impact of an individual being identified (which is not always a barrier).

What are the benefits of open data for policy?

There are three layers of benefits…

Indirect benefits

- New **business models**
- Broader **economic growth**
- **New purposes** for data
- Creativity and **engagement**

Direct benefits

- **Efficiencies and savings**
- **Collaboration** across public sector and with private and research sectors
- **Improved services**

Inside the org

- Improved **data quality**
- Greater **visibility** of data offerings
- **Awareness of data users**
- **Culture change** enabler

"Once released, other data scientists or 'data wranglers' can do the interpretation for you, **supporting more informed public debate**"

"If data is often released under the Official Information Act, openly releasing it **saves costs**"

"Agencies often duplicate data collection and info management systems, then don't share it (or do with multiple licenses) – open data **mitigates 'data hugging' and duplication**"

Open data is also a policy 'tool' and can support evidence-based policy advice – **see next two slides…**

Using open data as a policy 'tool'

Open data can be a policy 'tool' for:

➢ improving consumer experience

➢ more efficient delivery of interventions

➢ market interventions

➢ better data management interventions

➢ better policy making interventions.

Participants used design-patterns to explore how open data can support their policy objectives, in their particular contexts.

"Making things visual and tactile can help creativity"

Example: one table discussed "how to improve social licence for data collection and analytics?"

People are suspicious of what you are or might be doing

Building trust   (BETA)

When what you are doing is obfuscated, people can leap to conclusions about what's happening and your motivations. This can undermine trust, increase tension, distract effort and ultimately make it harder to achieve your goals. Being transparent and making open data available about what you're doing, particularly when coupled with accountability mechanisms, can help build trust.

Example: In the UK, people were suspicious of how the National Pupil Database (a closed dataset) was being used. By openly releasing requests to access it, public trust improved through transparency on how the data was being used.

Opportunities for data-informed policy advice

Looking ahead, there are both opportunities and challenges for open data to support evidence-informed policy advice

New combinations of data (open and closed) can create new knowledge and insights, and lead to whole new fields of application – think about which data sets maintained by others might be useful in your policy domains.

To date, much of the focus of open data initiatives around the world has been on the supply-side. For open data to be more responsive and sustainable, more insight needs to be gained on the demand-side and/or user needs.

"Maintaining data can be expensive, so ensure it generates knowledge needed for important policy issues – some data, when made open, may be more valuable for future than current governments; **such public assets are worth investing in given our stewardship obligations**"

'Big Data' technology can enable:

➢ linking together fragments of related yet heterogeneous information to identify undiscovered patterns/correlations and causations, including to support common-sense experience or test received wisdom

➢ the application of predictive analytics, which can increase the quality of scenario planning, including for better strategy setting and policy advice.

Barriers and enablers for open data

The biggest barriers for open data use are:

➢ practical (e.g. cost, unknown demand/value)

➢ social (e.g. culture of risk aversion, 'data hugging')

➢ technical (e.g. silos, system interoperability)

➢ legal (e.g. legislative restrictions).

> "Things can go wrong when individuals' privacy is breached – **we need to get better permissions and get better at de-identification, rather than stop trying**"
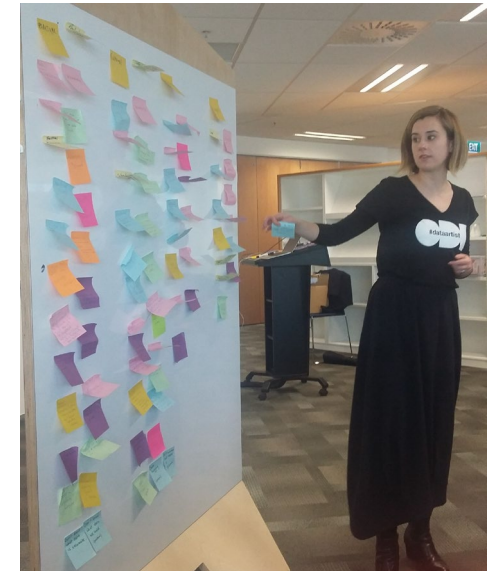
Practical and social barriers are the easiest to address, including through good meta-data practice and engaging communities of practice.

Practical enablers:

Useable period described

Availability period described

Discoverable from home page

Listed in a collection

Referenced from publication/application

Quality problems listed

Quality control process described

Social enablers:

Support for improving/fixing

Email support

Discussion groups/forums

Social media channels

Supported community

Tools and guides available to work with data

Conclusion and next steps

Key points on using open data for policy:

Open government data

1. Must be accessible, machine-readable, licenced for free use/reuse, non-personal, unclassified and non-confidential.

2. Should generate knowledge needed for important policy issues (current and future) and be linked to demand/user needs.

3. Supports collaboration and more informed public debate, mitigates 'data hugging' and duplication, and saves costs.

4. Can be used both as a 'tool' for different policy objectives, and as an input (evidence source) for quality policy advice.

5. Can be connected to other data, including through 'big data' technology, to reveal new insights, and improve knowledge and advice.

6. Can be enabled through practical and social means: good meta-data practice and engaging communities of practice.

Want to know more?

Any queries, feedback or help needed, please email: opendata@stats.govt.nz

If you want to learn more about open data, a range of resources including short interactive e-learning modules by the Open Data Institute is available here: www.data.govt.nz/toolkit/learn-more-about-open-data

You can follow @OpenDataNZ on Twitter for news, events, and good local and global examples.

Checkout GovHack projects, some of which are policy related: www.2017.hackerspace.govhack.org

Acknowledgements

Thanks to Ellen Broad and all participants for their constructive engagement, energy and insights.

If you have used open data for evidence-based policy advice, please email policy.project@dpmc.govt.nz and let us know so we can help Stats NZ demonstrate the value of open data.