

The 9Rs Framework:

A Worksheet for Establishing the
Business Case for Data Collaboration
and Re-Using Data in the Public Interest

November 30th, 2022

Lisa Talia Moretti, Andrew Zahuranec,
and Stefaan G Verhulst



Open.Data.Policy.Lab



THEGOVLAB

Introduction

Providing access to data can be a transformative act. As we have documented elsewhere, when access to data increases, its responsible re-use can improve governance and public services; empower citizens; create economic opportunity and accelerate innovation; and make science better and solve public problems ([Verhulst and Young, 2017](#)).

Despite these well-documented economic and societal value propositions, however, there has been limited access to tools and frameworks that can help organizations evaluate the business case of data access and re-use.

Why should an organization open its datasets? How does it support their organizational goals? What convinces leadership to invest in resources to make data available to others? In short, how to develop the business case for data openness?

Answering these questions can be hard without a supporting framework in place. Toward that end, we have developed the below worksheet, built around the “9R”s Framework, an evidence-based model for identifying the business case why an organization might open their data.

We help data stewards identify what each “R” entails (by listing a definition, inspirational case, and other considerations) and then provide them space both to indicate how important each “R” is for their organization and how it manifests. This evaluation exercise supports the creation of a compelling business case which could produce the kinds of insights needed to advance science and make decisions that create social value.



The “9R”s are divided into three basic kinds of return on investment and include:

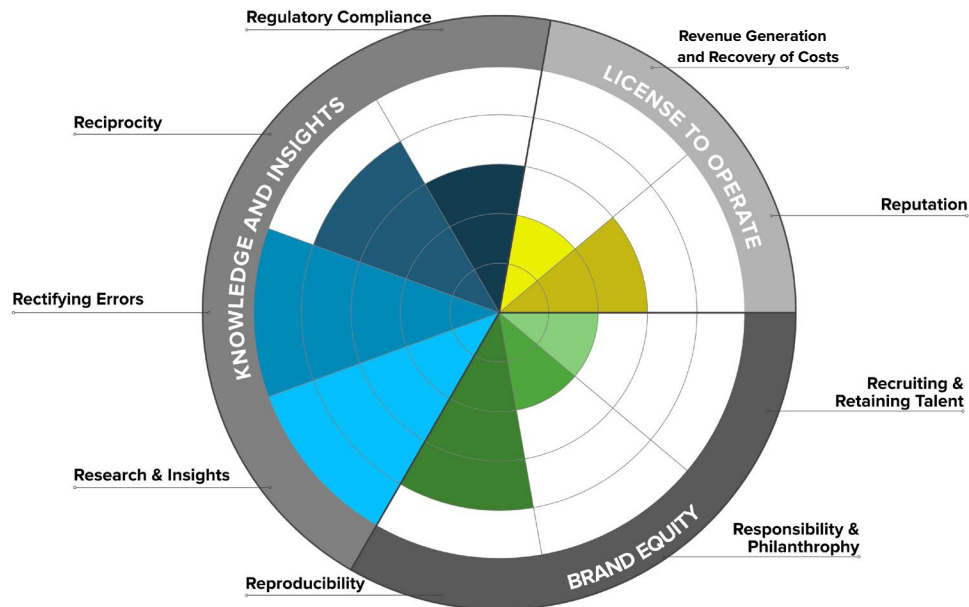
- **ROI Case #1: Increasing Knowledge and Insights.** The ways that data reuse can create a return on investment by generating new insights that an organization would otherwise be unable to extract from its information if it kept its dataset closed and inaccessible to outside parties. This case can include:
 - *Reciprocal exchanges;*
 - *Rectifying errors and improving data quality;*
 - *Research and insights;* and
 - *Reproducibility.*
- **ROI Case #2: Improving Brand Equity.** The ways that data reuse can create a return on investment by promoting an organization’s image to internal and external stakeholders. This case can be about improving an organization’s
 - *Reputation;*
 - *Responsibility and philanthropy;* and
 - *Retaining and recruiting talent.*
- **ROI Case #3: Establishing License to Operate.** The ways data reuse can create a return on investment by strengthening the mission or permission for an organization to operate. This case can be about issues of
 - *Regulatory compliance;* and
 - *Revenue generation.*

By examining the questions in the worksheet, data stewards can evaluate and map the benefits or motivations and consider which provides the strongest business case for providing access (if any).

The 9Rs framework also helps data stewards to assess solicitations for accessing data not only in terms of societal impact but also in terms of congruence with an organization’s overall purpose and values.

While completing the worksheet, it might be helpful to use a visual aid such as a spider chart to map the evaluation of each R. Below is an example of what a completed spider chart could look like once evaluation is complete. The figure below shows that generating new knowledge and insights offers the strongest motivation and relevance for this hypothetical organization, while brand equity and license-to-operate arguments are less relevant. We encourage you to use this or similar methods to evaluate motivations for reuse for your organization. A blank version of the below graphic can be found [here](#).

BUSINESS CASE FOR DATA REUSE IN THE PUBLIC INTEREST



 Open.Data.Policy.Lab

While the need for developing a business case is higher among private sector data holders, this document can be used by actors across domains— including public sector, civil society groups, and academic organizations. It relies on our ever-expanding research into opening up data, expert consultations, and industry interviews to explain what each “R” is, entails, and considerations involved in it.

If this resource was useful to you, please let us know at datastewards@thegovlab.org; and spread the word widely.

ROI Case #1: Knowledge and Insights: The ways that data reuse can create a return on investment by generating new insights that an organization would otherwise be unable to extract from its information if it kept its dataset closed and inaccessible to outside parties.

R1: Reciprocity

Definition: Reciprocity refers to what occurs when organizations that open their data gain access to other data sources and assets (e.g. expertise, technology) held by other institutions that could be important to business decisions. [Huston, Edge and Bernier \(2019\)](#) document how the real value of open data is not just about making datasets available for use. Often, opening datasets provides impetus for others to do so as well, creating a reciprocal data ecosystem and providing a multiplier effect that benefits all parties ([Apheris, 2022](#)). It can also lead to opportunities for data fusion in which one holder's dataset is enhanced by another, either by improving its accuracy or applicability for different use cases ([Chatzichristos, et al. 2022](#)).

Considerations: Organizations might want to be cognizant of the broader context in which their data is being used and how "reciprocity" might allow others to make unwanted insights. For example, they might think critically about whether reciprocal exchanges could contribute to the "mosaic effect" in which compilations of disparate datasets can be combined to reveal new sensitive insights that might undermine the privacy and wellbeing of individuals and communities ([McInerney, 2020](#)).

Ask yourself: By providing more access to its data, can your organization answer questions it might not have otherwise conceived or extract insights it might have otherwise lacked the resources, capabilities, or time to do so itself?



Very Irrelevant



Somewhat Irrelevant



Somewhat Relevant



Very Relevant



Most Relevant of All

Capture Notes Here:

Inspiration:

Consider an example from the environmental sector: Trusted data exchanges are crucial to monitoring the changes brought about by the climate emergency and to inspire innovative solutions to mitigate impact. Toward that end, The National Biodiversity Network Trust is a collaborative partnership created to exchange biodiversity information ([National Biodiversity Network, 2022](#)). With over 200 members including commercial organizations, supporters and sponsors, data is collected and shared within the network. The exchange works due to reciprocity: organizations know that by participating and providing data, they expand the data available to all, including themselves.

Additional Relevant Resources:

- [Collaborative Data Ecosystems](#): Article discussing how organizations that collaborate on data, machine learning and data science could be the key to solving complex problems.
- [Global Fishing Watch](#): GovLab case study on a data collaboration among SkyTruth, Oceana, and Google to map and measure fishing activity worldwide using data.

R2: Rectifying Errors and Improving Data Quality

Definition: A benefit of data openness is that it can improve the quality of data. When data can be accessed through open channels, it is typically reviewed and re-used by people other than those who collected and originally processed it who might be able to provide a fresh perspective or a different domain expertise. This review process can help identify errors and inaccuracies in the data that might not have been otherwise caught. The reviews and feedback from others can also help to improve the quality of future data collections.

Data quality can have a tangible impact on organizations and how they operate. Gartner has found that poor quality data is responsible for an estimated USD 12.9 million annually in losses ([Moore, 2018](#)). Writing for the EQUATOR Network, an international initiative to promote transparency in health research, Professor Dorothy Bishop explains that providing access to data “exposes researchers to the risk of being found out to be sloppy or inaccurate.” As with peer review, there is a strength in numbers to improve a dataset ([Bishop, 2014](#)).

Considerations: To help others review efficiently, organizations may wish to publish data collection methods, privacy statements, frequently asked questions, policies, and other information that provides context to those looking to re-use the data on offer. Providing a contact form so that people can ask questions or provide feedback may also prove helpful.

Ask yourself: By providing more access to its data or models, can your organization identify and rectify errors that might otherwise go unnoticed or unaddressed?



Very Irrelevant



Somewhat Irrelevant



Somewhat Relevant



Very Relevant



Most Relevant of All

Capture Notes Here:

Inspiration:

Open data's value in improving quality can be seen in the response to COVID-19. As the public and private sector both entered crisis over the disease, companies looked to data collaborations that could increase the quality of data available to public health authorities. Telecommunications data allowed leaders in the European Union to supplement and increase the granularity and quality of traditional health surveys that might predict where and how COVID-19 spreads. As The Atlantic magazine reports, COVID-19 has shown the need to “check multiple data sources and [...] ‘triangulate’ between them” while bringing an end to the “fragmented data status quo” that might lead to low-quality estimates ([Ladyzhets, 2022](#)).

Additional Relevant Resources:

- [Why We Should Care About Bad Data](#): A blog post by The GovLab that looks at the impact bad quality data has on decision-making
- [What makes quality open data](#): An online learning resource that covers data usability, standards and marques of quality as a researcher.

R3. Research and Insights

Definition: Increased access to data can lead to new research and insights that would not have otherwise been possible. By providing access to data, organizations can help advance the state of knowledge in their field. This, in turn, can help generate new questions, provide answers to old questions and lead to the development of innovative products and services.

Considerations: A major challenge for open data and data collaboration is ensuring they have sufficient data to inform meaningful innovation work. As one survey of open data projects notes, many initiatives are too technically focused and are “little more than websites linked to miscellaneous data files, with no attention to the usability, quality of the content, or consequences of its use” (Helbig, et al, 2014). Many organizations rely on “low-hanging fruit” that are easy to release but of little utility (Kitchin, 2013). The result is that data repositories become “data dumps” that cannot result in meaningful insights or action. This problem is closely related to the research–policy gap in which significant research effort is expended to generate insights but these insights are not acted upon or adopted (Martin, et al. 2019). It is often beneficial for organizations to have their work directed toward specific, tangible action. This problem is closely related to the research–policy gap in which significant research effort is expended to generate insights but these insights are not acted upon or adopted (Martin, et al. 2019). It is often beneficial for organizations to have their work directed toward specific, tangible action.

Ask yourself: By providing more access to its data, can your organization answer questions it might not have otherwise conceived or extract insights it might have otherwise lacked the resources, capabilities, or time to do so itself?



Very Irrelevant



Somewhat Irrelevant



Somewhat Relevant



Very Relevant



Most Relevant of All

Capture Notes Here:

Inspiration:

A desire for new insights that can yield solutions has long been a motivating factor behind open data portals. Swisscom, for example, used its reach across Switzerland to create an open data portal ([Swisscom Open Data, 2022](#)). The 44 anonymized datasets stretch across communication, mobility, network (mobile, railway and Swisscom hotspots) and sites (Swisscom shops). Created through the belief that innovation can be ignited through high-quality data, Swisscom shares data so that fresh, new thinking can be applied across Switzerland.

Additional Relevant Resources:

- [#Data4COVID19 Review](#): Report from The GovLab, with support from the Knight Foundation, reviewing lessons from the use of non-traditional data during COVID-19
- [The YODA Project Partnership with Johnson & Johnson](#): Short case study by The GovLab on YODA

R4. Reproducibility

Definition: When data is shared, it becomes possible for others to check and verify the results of the analysis. This ability to test results helps build trust and advance innovation agendas. For example, by making data available to others, organizations allow others to conduct identical or related work by providing an opportunity to confirm insights or revisit and improve upon previous analyses. In addition, by being open about data policies and processes, organizations can receive input and feedback into these and improve their data science capabilities as a whole.

Considerations: Amid the ongoing replication crisis, whereby shoddy findings get circulated because of a lack of external validation or journal due diligence, replication has proven especially persuasive for academia and a variety of other sectors ([Piper, 2020](#)). When trying to promote reproducibility, organizations may want to ensure their work focuses on real, quantifiable metrics, that they promote transparency in their methods as well as their data, and that they regularly verify their own processes internally to identify and counteract issues before they arise.

Ask yourself: By providing more access to its data, can your organization verify and validate its analysis, practices, and overall capabilities?



Very Irrelevant



Somewhat Irrelevant



Somewhat Relevant



Very Relevant



Most Relevant of All

Capture Notes Here:

Inspiration:

Consider the example of Microsoft Research's Open Data Repository. The repository, in Microsoft's own words, "makes available datasets that researchers at Microsoft have created and published in conjunction with their research." Run on its Azure platform, this repository provides a simple platform for Microsoft researchers and collaborators to share the basis of their work explicitly so that it can be validated by peers both within Microsoft's research community and those outside it. The platform gives internal and external researchers easy access to data, encouraging researchers to test one another's work and findings.

Additional Relevant Resources:

- [Identifying and addressing data asymmetries so as to enable \(better\) science](#): Article which seeks to start filling the analytical gap regarding global data asymmetries.
- [What the drive for open science data can learn from the evolving history of open government data](#): Article that looks at the history of the open government data movement and how the open science movement can identify ways to move forward and learn from it.
- [A manifesto for reproducible science](#): Paper, published in Nature, that argues for the adoption of measures to optimize key elements of the scientific process.
- [Should you share your data?](#): Blog post discussing the opportunities and challenges of sharing data as a researcher.

ROI Case #2: Brand Equity: The ways that data reuse can create a return on investment by promoting an organization's image to internal and external stakeholders.

R5. Reputation

Definition: One way in which providing access to data can help build brand equity is by enhancing an organization's image and reputation. When organizations open their data, it demonstrates their commitment to transparency and social responsibility. This, in turn, can improve the public's perception of the organization and build trust, which can attract new users and investors who value socially responsible organizations. The Mastercard Center for Inclusive Growth, for example, noted that it was important that they "show up as more than just a payments company" and demonstrate to potential partners that their company data "can be used for good."

Considerations: To turn an open data initiative into a brand-building exercise, organizations may want to create a compelling narrative that can be packaged into a variety of media formats and shared. Ideas could include building bespoke open data portal or creating branded content that can be shared across marketing channels.

Ask yourself: By providing more access to its data, can your organization improve its image and reputation to media, customers and/or investors who value socially conscious corporate actors?

Very Irrelevant
 Somewhat Irrelevant
 Somewhat Relevant
 Very Relevant
 Most Relevant of All

Capture Notes Here:

Inspiration:

The reputational benefits of data openness are especially important for organizations in the business sector, where data sharing can establish trust and help attract more customers at a lower cost and increase customer retention. The Open Data Institute quotes Sarah Hitchcock, partner at the location-analytics company GeoLytix, as stating that "open data products allow us to build a reputation in our sector and expand our customer base, but also contribute back to the ecosystem and help improve the quality of open data" ([Shadbolt, 2015](#)). Similarly, as Smita Jain, of the Mastercard Center for Inclusive Growth noted in an interview about her company's work with data sharing: "We want to show that we're committed to communities, committed to their geographies we operate in and it's not just a purely business transaction. We are there to be a supportive player in the community, too."

Additional Relevant Resources:

- [Data Responsibility: Using corporate data to save lives](#): TEDxMidAtlantic by The GovLab co-founder Stefaan Verhulst
- [Uber Movement](#): Explore how Uber is playing a role in helping cities grow in a way that works for everyone.
- Mastercard's [Data Responsibility Principles](#): Read more about their open data initiative, Inclusive Growth Score, here.

R6. Responsibility and Philanthropy

Definition: Data openness can also help bolster an organization's reputation as a responsible corporate actor. When data is shared, it can demonstrate an organization's commitment to generating social value. This can help attract new users, customers, and investors who value socially conscious corporate actors. Importantly, philanthropy does not need to be something that comes at the expense of other activities. Civic technologist Matt Stempeck, writing in the Harvard Business Review, notes that data philanthropy "allows companies to give back in a way that produces meaningful impact, and reflects the businesses' core competencies while preserving or expanding value for shareholders" ([Stempeck, 2014](#)).

Considerations: Organizations might want to think about what data releases can actually accomplish in what context. A well-defined goal of how data can be used for good—and by which audiences—can minimize the risk of publishing data for just publishing sake. Superficial data exchanges can be seen as "open washing"—superficial efforts to publish data without full integration with transparency commitments ([Verhulst and Zahuranec, 2022](#)).

Ask yourself: By providing more open access to its data, can your organization fulfill its social responsibilities, improving where they operate or achieving some social good?



Very Irrelevant



Somewhat Irrelevant



Somewhat Relevant



Very Relevant



Most Relevant of All

Capture Notes Here:

Inspiration:

Many companies report pursuing data openness out of a sense of obligation or social responsibility. Matt Webb, Head of Enterprise Data Management at UK Power Networks cites his organization's reasons for publishing portions of its data:

"The main driver is the recognition of the key role that data plays in the smart grid of the future and supporting the whole system interoperability, contributing towards Net Zero targets. We have a role to play as an organization; certainly from the perspective of corporate responsibility, but equally, if not more so in terms of our social responsibility."

This attitude matches with opinions in the United Kingdom, where 75% of adults are concerned about climate change ([UK Office of National Statistics, 2021](#)).

Additional Relevant Resources:

- [The Open Data for Social Impact Framework](#): A tool, developed by Microsoft, that organizational leaders can use to further understand how best to put data to work to solve important societal challenges.
- [Corporate Social Responsibility for the Data Age](#): Article that defines data responsibility, its three pillars and the cultural shift that is needed.
- [UK Power Networks' Open Data Portal](#): Visit one of the UK's biggest sets of data about the electricity network.
- [Sharing Data is a Form of Corporate Philanthropy](#): Article in the Harvard Business Review that argues that providing access to data should be considered a corporate social responsibility initiative.

R7. Retaining and Recruiting Talent

Definition: One of the most important ways in which data sharing can help companies is by allowing them to identify and attract individuals with the right skills and values. Data sharing also helps create a corporate culture that is attractive to top talent. When data is shared, it can help foster an environment of transparency and trust. This environment can be appealing to individuals who are looking for organizations that are responsible and accountable. As noted by Smita Jain, Manager, Impact Data Science of the Mastercard Center for Inclusive Growth, “Data for good really resonates with our employee base. It’s a very key point for those folks interested in working for Mastercard, and it’s good for the business because our talent is excited about working here and getting involved.”

Considerations: In their paper reviewing corporate social responsibility and organizational attractiveness, authors [Story, Castanheira and Hartig \(2016\)](#) find that “CSR can play an effective role in attracting potential employees.” The researchers note that “practices that improve the quality of the natural environment and the well-being of the society” are most convincing when combined with practices that “protect and develop” employees. Organizations might want to center these ideas if they pursue their own open data efforts.

Ask yourself: By providing more open access to its data, can your organization keep or attract diverse talent with projects that are compelling and socially relevant?



Very Irrelevant



Somewhat Irrelevant



Somewhat Relevant



Very Relevant



Most Relevant of All

Capture Notes Here:

Inspiration:

Google.org demonstrates these ideas can look in practice. The charitable arm of the Google company, Google.org harnesses the funding, tools, staff, data, and other resources of Google to support nonprofits, social enterprises, and civic entities around the world. While the fellowship has clear reputational benefits, it is also critical in keeping staff interested and engaged in data work, providing opportunities for employees to “make a difference” through data collaborations. In testimonies available on Google.org’s website, staff highlight how they were “motivated by the opportunity to combine technical skills” with their passion for other sectors and domains. Individuals also highlight their appreciation for working on topics that made “immediate and tangible difference in the community.” It encouraged them to remain with Google while allowing them to tackle issues they cared about.

Additional Relevant Resources:

- [Experimental Statistics](#): Article describing how the Office for National Statistics created an indicator of job vacancies with Adzuna job board data.
- [Google.org](#): Explore some of Google’s [open data sets here](#).

ROI Case #3: License to Operate: The ways data reuse can create a return on investment by strengthening the mission or permission for an organization to operate.

R8. Regulatory Compliance

Definition: As data has become a larger part of daily life and individuals have grown concerned about the concentration of data ownership, there has been growing demands for organizations to be more transparent about their operations. This in turn has been accompanied by increased attention to the need for organizations to open their datasets.

Many laws today require organizations to be more transparent and open. Government agencies in a number of jurisdictions are required to launch open data portals or publish data on their operations ([City of New York, NYC Open Data, n.d](#); [Kim, 2019](#)). Likewise, companies in the United States and other countries are increasingly expected to report on their financials, the management of their workforce, and other information ([Probst, 2021](#)). These efforts intend to boost transparency and the legitimacy of certain data efforts.

Considerations: Organizations who operate in multiple jurisdictions might want to look at the regulations and laws related to data collaboration where they operate. Rules, procedures, and expectations may vary from country to country and even within countries ([Klosowski, 2021](#)).

Ask yourself: By providing more open access to its data, can your organization better comply with regulations or improve its social license by demonstrating responsible data use and management?



Very Irrelevant



Somewhat Irrelevant



Somewhat Relevant



Very Relevant



Most Relevant of All

Capture Notes Here:

Inspiration:

For example, in April 2017, new legislation in the United Kingdom came into force that requires all organizations with 250 employees or more to annually report on their gender pay gap. This includes public authorities, private and voluntary employers. The gender pay gap is the difference in average earnings of men and women in the labor market. Created and published by the Government Equalities Office (GEO), gender pay gap reporting guidance outlines the data that organizations must collect, the methodology they must use to perform calculations and what kind of supporting narrative—an explanation of internal and external factors resulting in the gap—is expected. Organizations may also choose to publish strategies that describe how they intend to take action to narrow, and ultimately, close the gap.

Additional Relevant Resources:

- [Gender Pay Gap Service](#): Use the online service to explore data submitted about the gender pay gap.
- [State of Open Data Policy repository](#): Repository by the Open Data Policy Lab to assess policy developments on open data and data reuse.

R9. Revenue Generation

Definition: Private-sector organizations are incentivized to generate new revenue and cut costs. Providing access to data can help organizations accomplish both of these goals by providing opportunities to develop new revenue streams, and by reducing the costs through improved operational efficiencies. The Open Data Institute has published a guide on potential business models for open data—suggesting organizations look at freemium models, subsidize openness through the value it can provide other sectors, or rely on networks to support datasets ([Open Data Institute, 2013](#)).

Considerations: Data openness can present opportunities but organizations often struggle to scale up efforts due to inadequate funding among other issues ([Deiglmeier and Greco, 2018](#)). For efforts to succeed, institutions need to be willing to supply the start-up costs associated with them. They also need to be able to build broad-based support across their organization instead of limiting it to one department or team.

Ask yourself: By providing more open access to its data, can your organization find new opportunities to generate revenue or cut costs?



Very Irrelevant



Somewhat Irrelevant



Somewhat Relevant



Very Relevant



Most Relevant of All

Capture Notes Here:

Inspiration:

Cuebiq is a location-based measurement company that provides its customers with anonymized information on the locations of smartphone app users. One major branch of Cuebiq was its Data for Good initiative, which sought to provide “positive social impact through ethical and responsible use of location-based data” ([Cuebiq Marketing Team, 2019](#)). This program (now under the umbrella of Spectus.ai) had various motivations but necessarily thinks about financial sustainability. In the words of Cuebiq’s Brennan Lake:

“For shared value initiatives such as Cuebiq Data for Good to succeed in creating long-term social impact, it is important to integrate a financial sustainability component into the program’s overall strategy. We do not view Data for Good as a primary driver of profit within Cuebiq, but we do pursue a blended distribution model which includes both data philanthropy and double bottom line partnerships.”

Additional Relevant Resources:

- [Evaluating the Economic and Social Returns of Open Data Standards](#): A white paper from Frontier Economics evaluating the benefits of “openness” in different contexts.
- [Case study: The value of sharing data for improving market reach](#): A case study from the Open Data Institute on how a UK leisure operator shared data to better reach new customers.



Open.Data.Policy.Lab

opendatapolicylab.org