

The Periodic Table of Open Data: A User's Guide

Policies and Practices for Enabling the
Third Wave of Open Data

August 2022

Stefaan Verhulst and Andrew J. Zahuranec

Table of Contents

Introduction: 3

Periodic Table of Open Data Elements: 4

Diagnostic Tool: 5

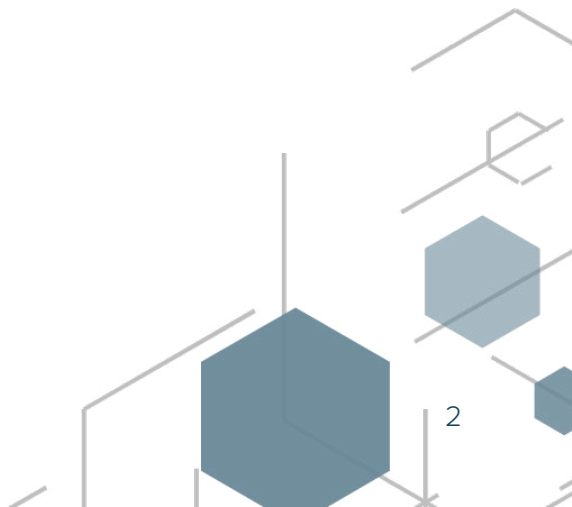
 Problem and Demand..... 5

 Capacity and Culture 7

 Governance and Strategy 10

 Personal and Partnerships 12

 Risk Mitigation 13



Introduction

Open data can be a valuable resource for society. When previously restricted data from the public, private, and academic sectors is made available for re-use by a broader range of actors, it fuels insights that would not have been possible otherwise, informing decisions on pressing social issues. This can empower citizens in making governments and businesses more accountable, enable disaster relief, and support people's livelihoods. In short, open data not only has normative value, but it can also lead to more innovative approaches to developing solutions to pressing public problems.

The question then is: *What makes open data “perform well”?*

At The Open Data Policy Lab (a collaboration between The GovLab and Microsoft), we not only seek to document the impact of open data initiatives around the world but we also try to identify the elements that lead to improved impact. This led to the exercise we undertook in 2016 when we sought to construct a “periodic table” that would list the enabling conditions and disabling factors that play an important role in determining the impact of open data initiatives. We were able to identify five central issue categories (as well as a number of sub-categories) that frequently play a key role in shaping the success or otherwise of open data projects across countries and sectors. Since then, we have validated the elements and complemented our research through numerous stakeholder interviews, workshops, and an on-going in-depth review of the literature.

Taken together these elements provide a conceptual framework to diagnose when initiatives are successful, and when they are not. Such assessments are always valuable in their own right, but they can serve a further purpose too. By reverse architecting assessments of success (and failure) across initiatives, we can identify a number of conditions that can be built into the design of open data initiatives, thus improving their chances of achieving meaningful positive impact. This document seeks to provide such a practitioner's guide.

The Periodic Table of Open Data: A User's Guide summarizes some of the major issues that open data practitioners may want to consider as they develop strategies for releasing and re-using open data in the communities they serve. In the below you will find:

- **The Periodic Table of Open Data (Version 3):** Depicting each of the elements of the Periodic Table of Open Data Elements and the categories in which they fit; and
- **The Periodic Table of Open Data Diagnostic:** A worksheet that readers can use to assess if an element is present in their project and, if so, how much weight it should be given in terms of importance for the success of the open data initiative.

We invite practitioners to use this guide to identify ways they can promote the success of their open data initiatives or mitigate risks.

Periodic Table of Open Data Diagnostic

This section of the document is a diagnostic tool intended to help data practitioners understand which elements are present in their open data initiative and the extent to which they need to address each of these elements in their open data initiative. Practitioners can use the sheet as a checklist and tool to weigh the importance of each element to their overall effort. The importance given to each element is its “atomic weight” (which is always context specific) and suggests how much priority it should be given.

At the end of this document is a brief section where readers can indicate the total “weights” of each category. Sections with higher weights are those that practitioners can focus additional attention on while sections with lower weights can be deprioritized.

Problem and Demand Definition

Problem and demand definition include all those elements that can allow an open data project to be targeted and optimized. They are about increasing the return on investment and minimizing wasted effort.

Benefits and Goals

Have you articulated a tangible value that the project will produce for its sponsors and the public?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Data Audit and Inventory

Has a data audit been conducted to understand which data is collected and generated and which can still be made available?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Data Ecosystem and Stakeholder Assessment

Has the broader sector and data ecosystem been assessed to identify the data, expertise, and other resources that could be made available via partnership?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Problem Refinement

Have you refined the problem to identify a specific purpose or issue to solve, the data needed to produce insights, and how those insights can be acted upon to produce real public value?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

User Research

Have you used user research to aid in identifying, mapping, and understanding a problem, its components, or user needs?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Capacity and Culture

Capacity and culture refers to those elements that relate to an organization's ability to pursue open data. It involves intangibles—including how individuals inside and outside the organization view the project—and tangibles—such as specific hardware and funding sources.

Culture and Institutional Buy-in

Have organizational leaders bought into and come to actively support the open data initiative?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Data Infrastructure

Is your data infrastructure mature enough to facilitate the easy collection, storing, managing, and sharing of open data?

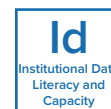


Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Institutional Data Literacy and Capacity

Have staff across the organization been trained on how they can use open data and data reuse to address the issues facing them?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Issue Salience

Has the open data project been connected to some issue of public or institutional interest that can be used to sustain active support?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Public Engagement

Have you promoted the open data initiative to the public in a way that conveys what it is and how it can meaningfully help them?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Resource Availability and Sustainability

Have you secured the internal and external sources of funding needed to systemize impactful and responsible open data?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Responsive Feedback Loops

Have you created mechanisms for data users and beneficiaries to provide input on an open data initiative?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Social License

Have you sought out the public's input and approval for the open data initiative which you are seeking to launch?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Strategic Leadership

Do you have an open data project leader who can enact policies, mobilize resources, and advocate for data openness?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Governance and Standards

Governance and standards are about how an organization makes decisions about the open data projects it manages and oversees. These elements relate to larger systems, structures, and organizing principles.

Data Licensing

Is there a robust data licensing regime that can protect and promote the re-use of data by outlining the conditions under which practitioners can use and re-use data?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Data Quality

Has the data been assessed for quality to allow individuals to derive meaningful insights from it?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Principles

Do the organizations working together on the open data effort have principles that align with one another and the work they are pursuing?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Data Standards

Have technical specifications—such as machine readability and data portability—for the datasets been laid out to allow it to be easily shared, exchanged, and combined with other assets?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Contracts and Data Sharing Agreements

Is there a contract or agreement that, in clear and specific terms, establishes what the roles and responsibilities of each party in the open data initiative are and how data will be made open?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Open by Default

Has data been made “open by default” to reach as many people as possible?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Policies, Positions, and Procedures

Are there policies, professional positions, and procedures that guide how individuals can best contribute to the open data initiative?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Personnel and Partnerships

These elements relate to the specific actors—both individuals and institutions—whose support can enable an open data initiative’s success.

Data Stewards

Are there data stewards present who are empowered to seek new ways to create public value through cross-sector data collaboration?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Domain Experts

Are there individuals with knowledge of a sector or domain who can be called upon to better identify opportunities to (re)use data for the public’s benefit or otherwise help identify, design, analyze, and communicate insights about the open data?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Third Party Supporters

Are there other organizations, institutions and individuals invested in a project’s success who might be useful in building a broad base of support for an open data effort?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Risk Mitigation

Risk mitigation elements are all those other practices adopted by an organization to prevent or mitigate potential harms that may arise in an open data initiative. These harms can be direct and obvious or relate to the more subtle ways in which data can exacerbate existing inequities.

Data Security

Are there practices in place to protect data against attack and breaches and promote public trust in their work?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Inclusive Design

Is the open data effort designed in a way that is mindful of inequities and seeks to reduce asymmetries in how impact is felt?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Institutional Alignment

Do all institutions working on the open data project share a common end goal?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Legal and Regulatory Requirements

Have the open data project's sponsors verified they are in compliance with all legal and regulatory requirements?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Privacy by Design

Are there procedures that have been put in place at the start of the data effort to preserve the privacy of data subjects and promote public confidence in the open data effort?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Proactive Data Holder Engagement

Are open data sponsors actively engaging with data holders to avoid missing or ignoring opportunities to use siloed data?



Importance

1	2	3	4
Not Important	Somewhat Unimportant	Somewhat Important	Very Important

Scoring

After answering each of the questions above, please tally your scores according to section. The highest scoring categories include those elements that are the most important for the success of the open data initiative while the lowest scoring categories are those that can be deprioritized.

Total Scores	
Problem and Demand	
Capacity and Culture	
Governance and Standards	
Personnel and Partnerships	
Risk Mitigation	



Open.Data.Policy.Lab

www.opendatapolicylab.org