

A Preliminary Process and Criteria for Evaluating Open Standards

Michael Roberts
Web Foundation
2015

Introduction	3
Background	3
Definition of an Open Standard	4
Preliminary Open Data Standard Initiatives Landscape	4
Controlled vocabularies	7
Process for evaluating open standards	8
Review criteria	9
Conclusions	10

Introduction

The Open Data for Development Programme facilitates strengthening coordination amongst open data initiatives to ensure they benefit citizens in developing countries. This includes improving the quantity and quality of open data supplied by governments, statistical offices, international development organizations and other relevant sources.

This report, supported by the Open Data for Development Programme, builds on that effort to define a preliminary process and criteria for evaluation of open standards for potential application and adoption by the International Open Data Charter Secretariat.

Background

The International Open Data Charter (ODC) includes six bold principles to guide the creation and implementation of open data policies and initiatives by governments. These include an expectation that all government data will be published openly by default, alongside principles to increase the quality, quantity, interoperability and re-use of the data that is released. In many cases this means publishing data using established open data standards.

In particular principle 4 on compatible and interoperable states:

- “1. We recognize that in order to be most effective and useful, data should be easy to compare within and between sectors, across geographic locations, and over time.
- 2. We recognize that data should be presented in structured and standardized formats to support interoperability, traceability, and effective reuse.
- 3. We will:
 - a. Implement consistent, open standards related to data formats, interoperability, structure, and common identifiers when collecting and publishing data;
 - b. Ensure that open datasets include consistent core metadata and are made available in human- and machine-readable formats;
 - c. Ensure that data is fully described, that all documentation accompanying data is written in clear, plain language, and that data users have sufficient information to understand the source, strengths, weaknesses, and analytical limitations of the data;
 - d. Engage with domestic and international standards bodies and other standard setting initiatives to encourage increased interoperability between existing international standards, support the creation of common, global data standards where they do not already exist, and ensure that any new data standards we create are, to the greatest extent possible, interoperable with existing standards; and
 - Map local standards and identifiers to emerging globally agreed standards and share the results publicly.”

Given the very clear requirement from the charter principles to publish data using open data standards, some guidance is therefore necessary to address how the Charter Secretariat

defines and in some cases, endorses a particular open data standard to leverage broad commonly used standards for increased compatible and interoperable data.

The charter itself provides a good base of criteria to be built upon later in this document for establishing how the published open data should be structured and the inclusiveness required amongst the community to ensure interoperability.

Definition of an Open Standard

According to the Open Data Handbook¹, an Open Standard is generally understood as technical standards which are free from licensing restrictions. This can also be interpreted to mean standards which are developed in a vendor-neutral manner.

Preliminary Open Data Standard Initiatives Landscape

The following outlines the Preliminary broad landscape of Open Data Standards Initiatives as they relate to international development efforts specifically. New open standards are emerging regularly in many sectors. Below is a preliminary landscape scan of open data standard initiatives. In many cases this work is ongoing and an established data standard has yet to be created. In the list below, the state of the initiative in the process of establishing an open data standard are indicated by early stage; draft stage; and implementation stage. Note this is not an indication of the maturity of the initiative but only relates to where in the process the initiative stands with regard to creating an open data standard. The early stage indicates the initiative has started but there is, as of yet, no draft open data standard. The draft stage indicates a preliminary data standard has been established and in some cases, some early pilots and use cases have been established. The implementation stage indicates the open data standard is now in use and is being actively revised by the community.

[Early Stage] **Grants and Contributions**

The Grants and Contributions standard is an emerging effort to describe domestic grant-making by governments. It is based on the IATI data standard and is an initiative of the Canadian Federal Government initially but designed to be broadly adopted globally.

[Early Stage] **Extractive Industries Transparency Initiative (EITI)**

The Extractive Industries Transparency Initiative (EITI) is a global Standard to promote open and accountable management of natural resources. It seeks to strengthen government and company systems, inform public debate, and enhance trust. In each implementing country it is supported by a coalition of governments, companies and civil society working together.

The adoption of the EITI Standard has seen a significant increase in the scope and detail of information being disclosed in EITI Reports. However, EITI data is often “locked” in pdf reports and is difficult to access, compare, and utilize.

¹ <http://opendatahandbook.org/glossary/en/>

The EITI Secretariat, in consultation with stakeholders drafted a framework for developing open EITI data standards for consideration in 2015. The report can be found here: https://eiti.org/files/scoping_study_for_eiti_data_standards_2015.pdf

[Early Stage] **Catalyzing Open Data for Extractives (CODEX)**

With support from Omidyar Network, the Natural Resource Governance Institute (NRGI) has launched a new initiative—"Catalyzing Open Data for Extractives," or CODEX—that will support efforts to go beyond mere transparency and ensure that available information is used. To maximize data utility, company and government disclosures are planned to conform to open data standards. NRGI is working to encourage harmonized and standardized reporting wherever feasible.

[Early Stage] **Open Budget Data**

The International Budget Partnership, along with a wide range of actors and institutions involved in public financial management, formed the multistakeholder Global Initiative for Fiscal Transparency (GIFT) to develop and establish international norms for budget transparency and participation. In addition to establishing clear global standards, GIFT is also focused on mobilizing the incentives and opportunities for governments to meet these standards.

Preliminary efforts to develop a data standard are ongoing and a preliminary effort to map the landscape of open budget data can be found here: <http://www.fiscaltransparency.net/resourcesfiles/files/20150902128.pdf>

[Draft Stage] **Fiscal Data Package - <http://fiscal.dataprotocols.org>**

Fiscal Data Package is a simple, open technical specification for government budget and spending data. It is a lightweight and user-oriented specification which aims to be extremely easy to use both for those publishing data (e.g. governments) and for those wanting to use the data (such as researchers and journalists).

[Varies - Draft Stage] **Mandatory Disclosures - <http://www.publishwhatyoupay.org/>**

Whether it is in royalties, bonuses or fees, extractive companies make significant payments to resource-rich countries around the world for their natural resources. Until recently, these numbers were shrouded in secrecy, which meant that citizens had no way of knowing how much money their governments were making from their natural resources, nor how those revenues were being spent. Publish What you Pay has been advocating for legislation that will oblige extractive companies to publish their payments per project in countries where they operate around the world. Currently a common template of common fields as a spreadsheet in the excel format has been created and adopted in Canada. The UK, and EU have also developed similar approaches.

[Implementation Stage] **International Aid Transparency Initiative Standard - <http://iatistandard.org>**

"IATI is a voluntary, multi-stakeholder initiative that seeks to improve the transparency of aid, development, and humanitarian resources in order to increase their effectiveness in tackling

poverty. IATI brings together donor and recipient countries, civil society organizations, and other experts in aid information who are committed to working together to increase the transparency and openness of aid.

At the centre of IATI is the IATI Standard, a format and framework for publishing data on development cooperation activities, intended to be used by all organizations in development, including government donors, private sector organizations, and national and international NGOs. It was designed in close consultation with key users of development cooperation data in developing countries, to ensure its relevance and utility for a variety of different data users.

Organizations implement IATI by publishing their aid information in IATI's agreed electronic format (XML) – usually on their website – before linking it to the IATI Registry. The Registry acts as an online catalogue and index of links to all of the raw data published to the IATI Standard”
Ref: <http://www.aidtransparency.net/about#sthash.Lmgo1142.dpuf>

[Implementation Stage] **Open Contracting**

The Open Contracting Global Principles encourage publishers to proactively disclose information and data on all stages of public contracting.

The Open Contracting Data Standard (OCDS) has been developed to allow information providers to share as much of this information as possible as structured, re-usable, standardized data.

The OCDS is a core product of the Open Contracting Partnership. The 1.0 Release Candidate of a data specification, as part of the standard, was developed through a rapid prototyping project. The project was undertaken by the by the World Wide Web Foundation and World Bank, and supported by The Omidyar Network and the World Bank.

[Implementation Stage] **Humanitarian Exchange Language (HDX)**

HXL, is focused on operational data—data that helps people understand what's happening on the ground during a humanitarian crisis. HXL defines data as information that software can read and analyse and encourages contributions in any common data format. A tool called Humanitarian Data Exchange (HDX) supports tabular data in CSV and Microsoft Excel formats, and for geographic data in zipped shapefile, kml and geojson formats.

[Implementation Stage] **360 Giving**

The 360 Giving data standard provides a common way to share transparent and interoperable information on grant-making. Their publishing model involves offering a simple spreadsheet format for publishing and consuming data, backed up by a structured data model, developer-friendly JSON serialisation, and conversion tools. The open data standard describes the whole grant making process and supporting in-depth analysis of grants, grantees and beneficiaries.

[Implementation Stage] **Open311 - <http://www.open311.org/>**

Open311 refers to a standardized protocol for location-based collaborative issue-tracking. The most developed function of Open311 is to report and track non-emergency issues in public

spaces. Common issues include potholes, broken streetlights, garbage, vandalism, and other problems that compromise public spaces and infrastructure.

An open standard is a key part of Open311. The current Open311 API specification being implemented in cities like San Francisco can be found on the API page of the wiki.

Controlled vocabularies

In addition to Open Data Standards, there also exists many sector specific controlled vocabularies. For example in the case of Agriculture initiatives such as “AGROVOC (a controlled vocabulary covering areas such as food, nutrition, agriculture, fisheries, forestry, environment etc) and OpenAGRIS (a world-wide bibliographic data repository in the agricultural domain) are already opening up and enhancing access to open data and agricultural research information by means semantic standards related to Linked Open Data.”²

In the case of Humanitarian Aid, there exists a host of controlled vocabularies, from ISO geographic codes³ to vocabularies maintained by the UN Statistical office.

In the case of the International Aid Transparency Initiative (IATI), in addition to ISO codes, IATI uses controlled vocabularies maintained by the Organisation for Economic Co-operation and Development (OECD).

In all of the above examples, controlled vocabularies play as important role as the open data standard itself.

² ref: Data standards; making CGIAR data available and accessible - <http://www.cgiar.org/consortium-news/data-standards-making-cgiar-data-available-and-accessible-2/>

³ <http://www.iso.org/iso/home.html>

Process for evaluating open standards

The following outlines a preliminary process by which Open Data Standards could be evaluated by the Open Data Charter Secretariat to ensure potential recommendations for adoption meet the necessary requirements outlined in the Charter Principles.

1. Proposal to Technical Working Group⁴

This is the initial stage in the process by which an open standard is proposed to the Charter Secretariat to be considered for inclusion as a recommended open standard within the resource centre. It is during this stage that a case would be made for inclusion. This proposal made may come from the Charter Stewards community, the Joined Up Data Alliance, or from the Open Standards initiatives themselves.

2. Prioritization

In this stage, the Technical Working Group would review the proposed open standards for inclusion within the resource centre and make a recommendation based on set criteria.

3. The working group recommendations would then be reviewed publicly for feedback.

4. Document - Before any final recommendation, a set of clearly maintained documents outlining clearly how to implement the open data standard for reference within the Open Data Charter resource centre would need to be established.

5. Final Acceptance by the Technical Working Group

At this stage the technical would ensure all the necessary requirements have been met, reflecting on the public feedback and ensuring documentation is available for governments implementing the data standard.

6. Announcement by the Charter Secretariat

Lastly at this stage a public announcement would be made by the Charter Secretariat.

The above process may include a yearly review of all existing open standards which in some cases may result in the open standard being removed as an endorsed data standard should the standard cease to be maintained and/or used by the Open Data community.

⁴ The Joined up data alliance could be a key instrument by which the Charter leverages technical coordination by those seeking to promote interoperability and compatibility amongst data standards. As such, this alliance plans to voluntarily take on common challenges faced by the Open Standards community where appropriate and to coordinate efforts with the Charter secretariat where possible.

Review criteria

The Open Data Charter itself clearly defines principles that would increase interoperability and compatibility between datasets as well as community engagement in the development process. In addition the Open Stand principles⁵ provide a solid foundation for the review criteria by which to determine endorsement of any open standard. The following outlines from general criteria to be used as general guidelines when reviewing an open standard.

1. Cooperation

Respectful cooperation between open standards initiatives, whereby each respects the autonomy, integrity, processes, and intellectual property rules of the others.

2. Adherence to Principles

Adherence to the five fundamental principles of standards development:

Due process. Decisions are made with equity and fairness among participants. No one party dominates or guides standards development. Standards processes are transparent and opportunities exist to appeal decisions. Processes for periodic standards review and updating are well defined.

Broad consensus. Processes allow for all views to be considered and addressed, such that agreement can be found across a range of interests.

Transparency. Standards initiatives provide advance public notice of proposed standards development activities, the scope of work to be undertaken, and conditions for participation. Easily accessible records of decisions and the materials used in reaching those decisions are provided. Public comment periods are provided before final standards approval and adoption.

Balance. Standards activities are not exclusively dominated by any particular person, company or interest group.

Openness. Standards processes are open to all interested and informed parties.

3. Collective Empowerment

Commitment by affirming open standard initiatives and their participants to collective empowerment by striving for standards that:

- are chosen and defined based on technical merit, as judged by the contributed expertise of each participant;
- provide global interoperability, scalability, stability, and resiliency;
- enable global competition;
- serve as building blocks for further innovation; and
- contribute to the creation of global communities, benefiting humanity.

⁵ <https://open-stand.org/about-us/principles/>

4. Availability

Standards specifications are made accessible to all for implementation and deployment. Affirming standards organizations have defined procedures to develop specifications that can be implemented under fair terms. Given market diversity, fair terms may vary from royalty-free to fair, reasonable, and non-discriminatory terms.

5. Voluntary Adoption

Standards are voluntarily adopted and success is determined by the community.

6. Technical

Adherence to the following technical criteria:

- The standard should represent data in structured and standardized formats.
- The data should be presented such they use common identifiers.
- The data standard is fully documented

7. Prioritization. What problem will this standard solve and what evidence exists demonstrating the open standard solves a problem.

Conclusions

This report has been designed as an input into a potential process for evaluating open standards for the application and adoption by the International Open Data Charter Secretariat.

By building upon existing globally accepted criteria such as the Open Stand Principles and processes that are inclusive, the opportunity exists to leverage the work of open standards initiatives through the Charter process, thus arriving at more compatible and structured datasets that can help citizens to fully benefit from the global open data efforts.