Open standards for data – user experience

A report presenting the findings of our user research into open standards for data

Open Data Institute

Contents

About

This report is part of a research programme funded by Innovate UK. The goal of the project is to provide guidance to organisations that need to find or develop open standards for data, so that they are able to do so more quickly and so that the standards they create achieve better adoption and wider impact.

The report has been researched and produced by the Open Data Institute, and published in February 2018. If you want to share feedback by email or would like to get in touch, contact us at research@theodi.org.

OPEN FOR FEEDBACK   How can it be improved? We welcome suggestions from the community in the comments.

Who is the intended audience for this research?

This report is for anyone interested in understanding people's experiences of open standards development and adoption.

The main purpose of this research is to inform the development of the ODI's own guidance and tools that support the creation of new standards.

We are publishing these results so that other organisations involved in open standards development can use the findings to inform and improve their own processes and tools.

We used a variety of methods to engage with a broad community of data users, data publishers and standards developers. These included:

- looking for existing user research on standards and data
- interviewing data experts about their problems, experiences and solutions
- Surveying data users to identify the types of problems they encountered when using data and data standards
- running problem mapping and co-design workshops with data experts and standards developers
- testing existing ODI guidance and documentation with new audiences to gather feedback

By engaging with users we tried to answer some broad questions:

- Who is developing open standards?
- Why are they developing open standards?
- What does a good open standard look like?
- How do people develop them?
- What are the barriers people encounter?
- What biases might impact our interpretation of the results?
- Have any of these questions already been answered?

What we found

Our findings cover various aspects of standards development and adoption, and shaped our work for the rest of the project

Each finding includes:

- a typical quote from a user
- a brief discussion of the finding
- a recommendation

The first two findings were identified early in the research process. They helped shape our work during the rest of the project. The rest of the findings cover various aspects of the process of developing and adopting standards, and reflect the variety of views and experiences of the users we talked to.

Avoid talking about "open data standards"

" *I don't know what you mean by 'open data standard': do you mean a standard for publishing open data?* – nearly everyone we spoke to

Not many people understand the phrase "open data standard". Is it a standard for open data? Or a data standard that is open?

Through conversation with people we settled on "open standards for data" as the best way of describing our area of interest: how do people publicly work together to agree how to organise and/or publish data? Whether the data is published as open data is a secondary concern.

Proposal: use the phrase "open standards for data" to describe openly available standards for the naming and organisation of data.

A successful standard is a useful standard

> *You can have the most perfect standard but if no-one uses it, it's pointless* – expert working on standards without a mandate to enforce adoption

To understand how different approaches to developing an open standard could affect its success, we tried to identify the characteristics of a successful open standard.

The number of people using an open standard was one common indicator of success.

Users suggested that a successful open standard:

- solves a particular need or problem
- works with other standards
- shows what decisions were made while developing the standard and why they were made
- has been implemented and tested
- is robust and as simple as possible
- is developed using a process that works for its community
- is developed using a process that delivers early results
- Is developed by people with a variety of skills and backgrounds

Users told us that developing a successful open standard includes:

- understanding the problem
- getting funding
- finding others to work with
- designing the standard
- supporting and maintaining the standard
- communicating the benefits of the standard

This work might be carried out in sequence, or might be carried out in an iterative process. There seem to be pros and cons for each approach.

People told us that having contributions from people with different skills and backgrounds provided a constructive tension during the development of an open standard. They believe it takes researchers, developers, managers, domain experts, business analysts and community engagement managers to create a successful open data standard.

Proposal: solve a small problem quickly to see if the standard is useful.

Finding existing standards is an unsolved problem

> *How do you find existing work on a data standard?* – standards developer

We did not focus on finding existing open standards, but it became clear that discovering an existing standard or relevant work in an area is an unsolved problem.

Data users don't tend to look for existing standards before they publish a dataset. Even for people who regularly work with standards, discovering standards relevant to a particular area is difficult.

Open standards experts appear to largely discover open standards through formal and informal networks. These include standards bodies, conferences, hierarchical teams or working groups, Twitter, blogs, open datasets, and GitHub projects.

Once they have found an open standard they need to determine if it is relevant. There are three immediate questions that all standards must answer in order to be useful:

- was this open standard designed to meet my needs?
- how do I use this open standard?
- is it technically possible for me to use this open standard?

It is often difficult to answer these questions because the examples and documentation are not accessible and the design decisions behind the standard are unclear.

This lack of accessible tooling is one of the reasons why people think that using an open standard will be almost as much work as creating one to meet their own needs.

They may also have to decide whether it is worth the time to contribute to a shared standard which is likely to be slower to deliver.

Better documentation and tools that are usable by a broad range of people are likely to reduce the cost of both adopting an existing standard and contributing to an emerging one.

Proposal: help potential users find and use your open standard by providing useful summaries of its goals, status and level of adoption.

Examples and metaphors help to explain the value of standards

> " *Reach out to people who don't believe data is their thing* – senior leader at a government agency

People use a number of metaphors to help explain the value of open standards to "non-technical" people. Metaphors and easy to visualise examples were used across organisations.

For example:

- **Rivers**: Almost everyone mentioned rivers (or roads) as a way of demonstrating the value of standards. Rivers flow between communities, regions and countries. There is a need for each region to have a shared understanding of what a river is so that it can be monitored and maintained consistently, regardless of geopolitical borders. Within each of these territories there are multiple organisations using or reporting on the river: wildlife conservationists, boat-hire companies and government organisations. If there's a problem with the river all of these organisations may need to be able to talk about that problem in a way that the other organisations can understand. This metaphor was used to show how different organisations and countries need to build relationships through standards in order to have consistency and certainty in the data

- **Acrobatics**: An end user has always wanted to have a go at acrobatics. They do a local search and find several classes. Some results are very specific: "double trapeze acrobatics", "rope acrobatics", "silk acrobatics" and "aerial cradle". Others are simply called "acrobatics". As the end user how do I know which one is right for me? The user might have to attend a few classes before realising that rope and silk acrobatics are the same. Or that one of the classes titled "acrobatics" is also a rope acrobatics class. This metaphor was used as an example of how an end user (consumer) need can shape or drive definitions of standards before they need to be widely adopted

- **Historical comparisons**: People with government backgrounds stressed how not using paper has changed the ways we collect, store and present data – we are no longer limited by paper. Others with industry backgrounds reflected on manufacturing and how increased accessibility to computers and spreadsheets has paradoxically devalued data as anyone can mindlessly collect, structure and present data. Others stressed how standardisation has always been essential to production lines. These examples highlighted how standards aren't new, but that as sharing and using data becomes more accessible, they will become more necessary and have the potential to be more powerful

Proposal: use metaphors to help explain complex concepts.

Names matter

" *...increasingly businesses are recognising that there are multiple ways of representing the same thing* – senior leader of government agency

Talking to each other is hard. When developing a standard many people mentioned the need to clearly understand what another person meant by a certain concept. For example: *"What do you mean when you say river? I think that's a beck because X, Y, Z"*.

An important part of developing an open standard is naming the relevant concepts and agreeing definitions that are not too restrictive.

However people who have not been involved in the creation of an open standard often question the meaning of these definitions with edge cases. It seems important to be able to see how and why these definitions were adopted.

Proposal: ensure that any terms and concepts you are using are in scope, are clearly defined and are properly documented.

Early translation encourages understanding

" *Translating is a good way to know what you want to say* – European government agency

Many people highlighted the need to work with people from different organisations who have different words for the same concept. This is particularly difficult when collaborating with people in other countries.

A number of people we spoke to were developing standards for use in an international context or needed to support users in multiple languages.

Lots of people spoke of the difficulties of translating lengthy documents that use complex concepts to make them accessible to a wider audience or to translate them into other languages.

Translation is generally perceived as a problem. Some people pay organisations to translate their work but this slows them down, and can make the cost of updating documentation high.

However some argued that translating your work early can help involve more users and reveal undetected assumptions and concepts that you need to translate. This can save you confusion later when there are multiple implementations of the standard or as concepts change.

Proposal: where necessary, translate specifications and guidance early.

Show your working

> **"** *Governance of change is really important –* standards developer

People have developed different approaches to show how they make decisions for different purposes:

- **models of processes or timeline** to show when and how people can contribute to the development of a standard. These vary between standards bodies and organisations

- **minutes of meetings or conversations** as a reference when people ask a question that has previously been answered

- **publishing a roadmap of what the standard will cover** as a way of focusing and managing contributions without distracting from the agreed immediate work

All of these approaches aim to provide provenance and context: how the standard has been designed to meet a set of needs and the methods it used.

Proposal: work in the open, using platforms that are designed to support collaboration.

Power, money and relationships are always a concern

> **"** *In practice it's built by [a monopoly] with little representation from publishers or software companies who use these standards* – director of a private media software company

Anxiety over politics, monopolies, and control permeated every conversation around standards.

There were several examples where experts had spent years crafting a standard only for Google to release a new standard that was a solution to a simplified version of the problem they were working on.

Most people felt there was more to be done to persuade senior management that the effort of contributing to a shared open standard was worth the cost.

People attempted to show how working on open standards was a good investment of time in a variety of ways.

They used different techniques depending on the relationship between the standards developers and the users of the standard. The less authority the standards developer had over potential contributors, the more they needed to focus on building relationships and persuading others to contribute. There were several examples of approaching this problem including:

- agreeing codes of practice between agencies
- running cross-organisation workshops
- referencing best practices elsewhere to influence behaviour
- mapping ecosystems to understand useful interventions
- developing model business cases to encourage adoption

The need to persuade was particularly important for smaller organisations who depend on funding cycles. They needed to be able to measure and describe the impact of their work in ways that clearly show the benefits to organisations that might co-create or adopt the standard.

Organisations also needed to be convinced of the benefits of committing themselves to an open process which might result in them having less overall control in the direction of the standard.

Proposal: involve and persuade partners and users, and clarify who owns and maintains a standard.

Teach through experience

> *There's often a gap between those who create standards and those who use them –* representative of a large standards body

Standards are often very long. Some standards are thousands of pages long and specific to a particular technology or area of expertise. As a result the documentation is seldom read.

People with different technical environments or expertise to the creators of the open standard can find it difficult to know how to use the standard. It requires significant time to investigate.

Teaching people how an open standard should be applied is rare. Some have mediated complex specifications by concentrating on the experience of using the open standard through blogging, developer documentation and tutorials.

Despite this there still appears to be a consistent knowledge gap between the people who have developed a standard and the people who might adopt it.

Proposal: recognise that specifications are reference documents and not guidance. Provide additional support to drive adoption.

Lightweight standards are easy to use

> *… do one thing well –* many people

Several people made the point that the more that a standard tries to do, the harder it is to develop – and the harder it is to understand and to apply. This slows both development and adoption.

Many people recalled creating a detailed technical standard, only for it to be overtaken by a lightweight schema that was viewed as being too simple, but which was quickly and easily adopted.

Others warned that standards were hard to change once they were being used and there was a need to get things right first time. Almost everyone agreed that standards needed to be adaptable and modular.

The Office of National Statistics ethnicity classifications was a popular example of how a smaller standard might be incorporated into other data standards. The ONS have done the hard work in standardising a difficult and sensitive area. And, in the UK, if you are collecting data about ethnicity, then at some point you are probably going to compare the data you have with census data about the UK population.

The standard is so well used that it can't be agile and respond to change without requiring the same agility of the people who use it.

Proposal: build small standards that can be reused and built on by others.

Consult or fail

" *We've struggled to know who to talk to –*
service manager of digital service at a Quango

Almost everyone told us their processes are "open". But it was clear that "open" means different things to different people. For example, one open process relied on inviting people to a face to face meeting in Switzerland. Another was described as open because the results were published online, where people could find it. Other processes were more robust: they involved many opportunities for participation both online and offline throughout the standard-making process.

Poorly designed consultation and governance can cause problems with standards development, leading to incorrect decisions or poorly timed feedback.

Developers often struggle to reach a diverse group of people. They can end up simply validating their preexisting beliefs on how an open standard should be designed.

Often, requests for feedback don't produce results until the very end of the development process, when there is a sudden rush to provide comments. This might involve dealing with last minute (hypothetical) edge cases, or having to deal with a variety of technical issues.

Making last minute changes is also difficult if it is not clear, up front, what level of concern or feedback should trigger redrafting or changes to a standard.

Proposal: draft the standard in the open, make your governance clear and put effort into ongoing engagement.

Summary of proposals

We have identified a number of ways in which standards owners and developers can improve their approach to scoping, developing and driving adoption of a data standard. These are:

- use the phrase "open standards for data" to describe openly available standards for the naming and organisation of data

- solve a small problem quickly and deliver quickly to see if the standard is useful

- help potential users find and use your standard by providing useful summaries of its goals, status and level of adoption

- use metaphors to help explain complex concepts

- ensure that any terms and concepts you are using are in scope, are clearly defined and are properly documented

- where necessary, translate specifications and guidance early

- work in the open, using platforms that are designed to support collaboration

- involve and persuade partners and users – and clarify who owns and maintains a standard

- recognise that specifications are reference documents and not guidance – provide additional support to drive adoption

- build small standards that can be reused and built on by others

- draft the standard in the open, make your governance clear and put effort into ongoing engagement

Our intention is to use this insight to help inform the development of guidance and tools to support standards developers. We also hope that this summary of our user research might be useful for others who may want to build on our work, to create additional tools and guidance, or to explore some of these issues in more detail.