# What makes quality open data

Assessing how usable open data is cannot be done quickly.

There are a number of community-based standards and quality marques that can help you assess how usable data is.

**In this module we’ll explore the following:**

What makes data usable

How standards help increase the usability of data

Marques of quality

# Assessing open data quality - in 2 minutes

ODI Trainer David Tarrant explains what quality data is, how to recognise it and which quality marques consumers of open data should look for.

**What is open data quality?**

Open data becomes usable when a human can understand it and a machine can manipulate it.

Users of open data need the permission of its publisher, granted by an open licence. But the open licence alone is not enough to guarantee the usability of data.

It is unlikely that people will be able to engage with data to derive insights and demonstrate benefits unless it is usable.

# Community-driven standards

Community-driven standards define the legal, practical, technical and social requirements for an open dataset to be useful.

## Legal requirements

In addition to the open licence, there are three legal requirements that need to be considered. You must:

Protect sensitive information like personal data.

Preserve the rights of data owners.

Promote correct use of the data.

## Practical requirements

It is important that open data can be found and relied upon by users.

This means that open data publishers must:

Link to the data from their website.

Update the data regularly if it changes.

Commit to continue to make the data available.

## Technical requirements

There are three recommendations that define the technical aspects of open data:

The format in which the data is published.

The structure of the data.

The channels through which the data is available.

## Social requirements

For data use to be sustainable, it is important to have an engaged community of users. The best datasets have:

Active support channels.

Discussion groups and forums.

Published how-to guides on working with the data.

# Marques of quality

Marques of quality provide users with confidence about the quality and usability of open data.

There are two key indicators of open data quality:

## 5 Stars

The '5 Stars of linked open data' include the basic legal definition and technical excellence of the published data.

The 5 Stars were developed by inventor of the Web, Sir Tim Berners-Lee, in 2001. The 5 Stars make recommendations for the format and structure of the data.

You can learn more about the 5 Stars in modules 10 and 13 of this course.

## ODI Open Data Certificates

The Open Data Institute's Open Data Certificates are practical, social, legal and technical guidelines to consider when publishing or using open data.

The certificates act as a self-assessment questionnaire for publishers, and in some cases users, to assess and benchmark the full usability of a dataset.

[Take me to the ODI Open Data Certificates site](https://certificates.theodi.org/)

**Are you ready for quality open data?**

Can you recall the practical, technical, legal and social requirements for an open dataset to be useful?

**What is quality open data?**

**Open data becomes usable when...**

it is imported into our system a human can understand it, machines can manipulate it and it has an open licence the government publishes it

**That’s right!**

Open data becomes usable when a human can understand it, machines can manipulate it and it is made available under an open licence.

**Are you sure?**

Open data becomes usable when a human can understand it, machines can manipulate it and it is made available under an open licence.

# Legal considerations

**There are three legal considerations as part of the standards for publishing open data, one of which**

**is...**

not to break the licensing law to publish any type of data to protect sensitive information such as personal data

**That’s right!**

In addition to the open licence, there are three legal requirements that need to be considered. Protect sensitive information like personal data, preserve the rights of data owners and promote correct use of the data.

**Are you sure?**

In addition to the open licence, there are three legal requirements that need to be considered. Protect sensitive information like personal data, preserve the rights of data owners and promote correct use of the data

# Technical considerations

**There are three technical considerations as part of the standards for publishing open data, one of which is...**

what format the open data is published in whether it can it be used on a computer whether it takes up much bandwidth

**That’s right!**

There are three recommendations that define the technical aspects of open data. The format in which the data is published in, the structure of the data and the channels through which the data is available.

**Are you sure?**

There are three recommendations that define the technical aspects of open data. The format in which the data is published in, the structure of the data and the channels through which the data is available.

# Marques of quality

**The Open Data Institute advocates the use of its own quality marque known as...**

the ODI big green tick the ODI stamp of approval

Open Data Certificates

**That’s right!**

The Open Data Institute has developed Open Data Certificates which include practical and social requirements alongside in-depth legal and technical guidelines to consider when publishing open data.

**Are you sure?**

The Open Data Institute has developed Open Data Certificates which include practical and social requirements alongside in-depth legal and technical guidelines to consider when publishing open data.

# Standards + marques = quality open data

Open data becomes usable when a human can understand it and a machine can manipulate it.

Both of these require the permission of the publisher, granted via an open licence. However, the open licence alone is not sufficient to guarantee the usability of data. Simply putting a file on a website with an open licence does not constitute open data.

It is unlikely that users will be able to properly engage with unusable data to derive insights and demonstrate benefits.

In the next module, we compare open data initiatives and look at the key performance indicators of successful ones.

Next module

Main menu

# Module 5 – video

I’m Dave Tarrant, I’m the Trainer and Data Scientist here at the open data institute. So open data becomes useable when it’s understandable by humans and can be manipulated by machines. The most important aspect, as we’ve mentioned is that it must be available under an open licence. But this is not sufficient to guarantee that data is useable. Simply putting a file, or dumping a file on a website with an open licence, does not guarantee the user will understand the context under which that data can be used. And so, to make open data useable we must strive for higher standards of quality. Fortunately, these standards are defined by the community, and they define the practical, legal, technical, and social aspects of high quality open data. So the community standards define the legal, practical, technical, and social requirements for higher quality useable data. So the legal side is all about protecting the sensitive information, or maybe the personal data, preserving the rights of the owners of that data, whilst also promoting the usefulness of that data and the usability. The practical side is all about the what, the where, what is that data, where did it come from, and how is it been used, providing the context. The technical side is about the formats that the data is available in the structures of that data, and the channels through which it is available. And finally the social side is all about building communities so it recommends support channels, how to set up and establish discussion groups and also tools and guides to work with the data. So there are two key indicators of quality open data. The first one is the five stars of linked open data which defines the legal and the technical excellence of the data. The second one is the open data certificates and this adds the practical and the social requirements whilst also exploring in more depth the legal and the technical requirements. So either of these marks on a dataset will be able to provide you with an instant assessment of its usability.