**Module 5 – Dave Tarrant – Accessing open data**

**Accessing open data**

I’m Dave Tarrant, I’m the Trainer and Data Scientist here at the Open Data institute.

**What determines open data quality?**

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.

**What makes high quality open data – practically, legally, technically and socially?**

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.

**How can I look for high quality datasets?**

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.