

Video Transcript Lesson 2.2 What is Open Data

I'm Dave Tarrant, I'm the trainer and data scientist here at the Open Data Institute. So open data is data that anyone can access, use and share. So data becomes accessible by being made available online, on the web.

Data becomes usable by being available in a common, understandable format. The most important thing is that data is licensed. This license should allow for free use of the data, including for commercial purposes.

So think of data at the bottom of a three-tier triangle. So you've got data, information and knowledge. So data is the lowest level of abstraction from which information and then knowledge are derived.

So think of all the locations, pictures, reviews, descriptions and prices that you might need in order to plan a holiday, for instance. So you take all of this data and you start turning it into information about where you want to go. And from that you contextualise it and turn it into knowledge that is relevant to you.

At the same time, someone else could take all that data and look at the differences in cost between activities across the country. So a completely different usage. So it's critical to have this data so we can build these different levels of information and then allow people to contextualise them to build their own knowledge.

So data is open if anyone is free to use, reuse and share it. So this use must be without restrictions. So if you apply a non-commercial license on data, that's a restriction.

So therefore it's no longer open. But although data is free to use, and that is how it's measured in its usage, it's all about use, that doesn't necessarily mean it comes without cost. It might not be free to access.

In order to maintain and sustain the availability of this open data, there may be some marginal cost. And the common occurrence is that this cost should be no more than the reasonable reproduction cost of the unit of data that someone is asking for. So something to be open data is all about measuring how people can use it.

It's not about how it's published and how it's made available. It's about what people can do with it once they have it. That's what makes it open data.

Open data has the power to transform the interaction between governments, businesses and society. And this will unlock real value. New applications can develop.

Problems can be solved. And this will help generate real value for society, the economy and also solve some of the environmental issues. And this is where it's already generated real impact in improving our transport infrastructures and saving countless lives.

The web is already an accepted part of our societal infrastructure, and open data is the next critical part of this infrastructure.