

# Ensuring Data Quality

David Tarrant  
@davetaz

# Definition of Open (OKF)



A piece of data or content is open if **anyone** is **free to use, reuse, and redistribute** it — subject only, at most, to the requirement to attribute and/or share-alike.

# Guidelines



5-Stars

# 5-Stars



<http://5stardata.info/>



OL RE OF



OL RE



OL



OL RE OF URI

<http://data...>



OL RE OF URI LD



# 5-Stars



Available on the web (whatever format) *but with an open licence, to be Open Data*



Available as machine-readable structured data (e.g. excel instead of image scan of a table)



as (2) plus non-proprietary format (e.g. CSV instead of excel)



use URIs to denote things and when people look up these URIs, provide useful information and/or data.



All the above, plus: Link your data to other people's data to provide context

# 5 Aspects

Contextual	
Legal 	Practical
Technical 	Social

# Open Data Certificate



<http://certificates.theodi.org>

# 5 Aspects

## Contextual



Legal



Practical



Technical



Social





# Adding Context

<p>Who</p> <p>Collected it? Owns it? Publishes it? Is the Audience?</p>	<p>Where</p> <p>Was it collected? Is it used? Is it described? Is it located?</p>
<p>What</p> <p>Is it (title/description)? Type of data is it? Type of objects?</p>	<p>When</p> <p>Collected? Published? Updated? Due next update?</p>

# General

				
Title/Description/Publisher/URL	•	•	•	•
Release Type	•	•	•	•

# Legal

Rights to publish ©

Data licensed 

Content licensed 








Clear privacy statement 

Sources of data documented

Audited anonymisation 



# Practical

				
Usable period described		•	•	•
Availability period described		•	•	•
Discoverable from home page			•	•
Listed in collection			•	•
Referenced from publication/application			•	•
Quality problems listed			•	•
Quality control process described				•

# Technical

Data hosted online



Type of data defined

Machine readable metadata

Clear technical documentation

Persistent & common identifiers used

5-Star Linked Data



Machine readable provenance

Data can be verified



•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

•

# Social

Support for improving/fixing

Email support

Discussion groups/forums

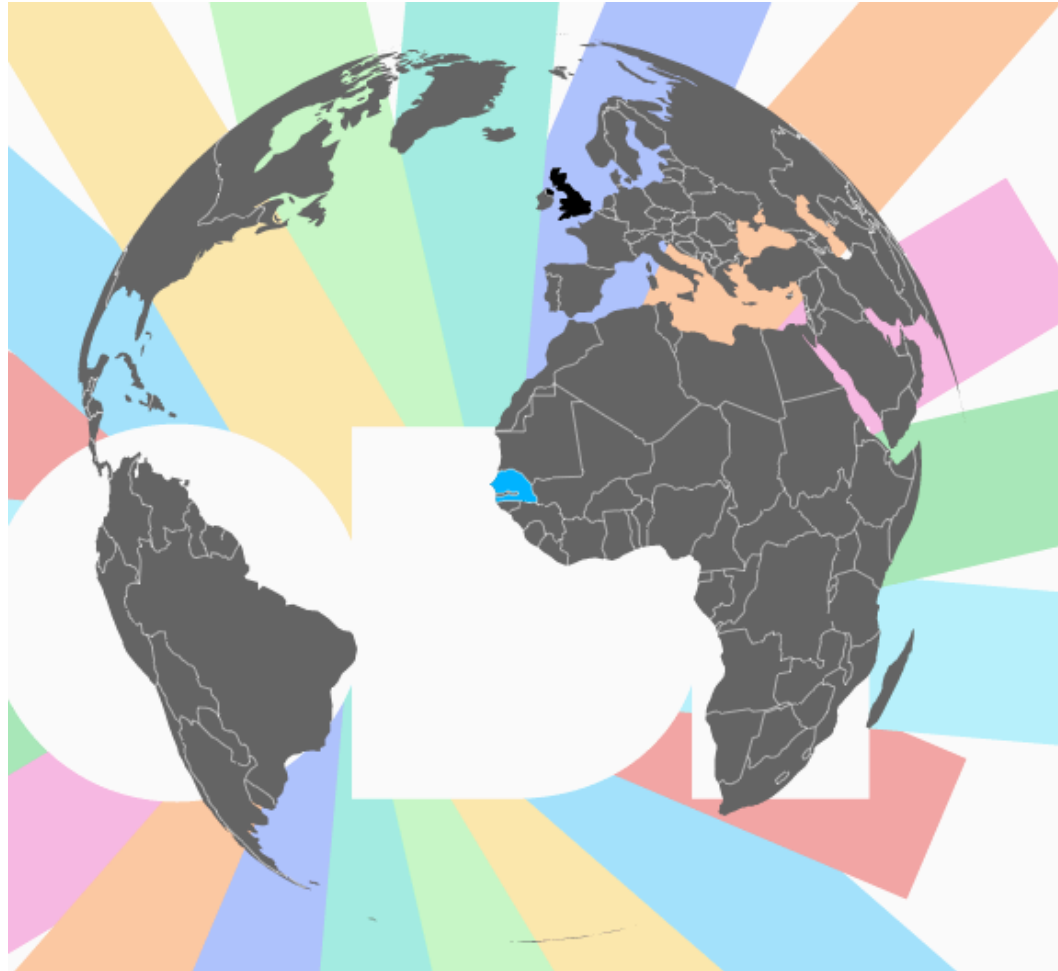
Social media channels

Supported community

Tools and guides available to work with data



# Open Data Certificate



## Where can you create certificates?

Different countries have different laws, so their certificates are slightly different too. Choose your country to create the right certificate.

Senegal

Create a certificate

**SN** alpha

This certificate hasn't been properly **localised** yet, and we need your help! Please contact [certificate@theodi.org](mailto:certificate@theodi.org)

### Legend

Certificate in alpha stage   Certificate in beta stage   Certificate in final stage

# One Objective

Minimum Requirement

3-Star Data

Standard Level Open Data Certificate







Thank you