

Ertugrul Bircan Copur

DATA QUALITY MODELLING AND MEASUREMENTS A case study with German open government data

Goals



Set of Data Quality Metrics



Survey and Participation

We asked application developers about: - Their experience in programming - Their experience with open data – The main and additional minor problems with the dataset – The main and additional positive aspects of the dataset Low participation – only four participants unusable for us)

Five datasets that we can evaluate (six in total, one

Evaluation by Metrics

A set of characteristics based on the SQuaRE-aligned Portal Data Quality Model (SPDQM). – Completeness, Traceability, Compliance, Accuracy, Understandability, Currentness, Expiration A total of fifteen metrics were used to evaluate the quality of

the datasets.

Results and Comparison

Dataset	Completeness		Traceability		Compliance		Accuracy		Understandability		Currentness		Expiration	
	Survey	Metrics	Survey	Metrics	Survey	Metrics	Survey	Metrics	Survey	Metrics	Survey	Metrics	Survey	Metrics
1 – Used Glass Containers	0	+	0				0	+	0	0	0	n/a	0	
2 - Popular First Names		++	0	+	0	++	0	+	0	0	0	n/a	0	+
3 - Christmas Markets		++	0	+	0	++			0	0		n,∕a	0	
4 – Bus Stops	+		0	+	0	++	+				0	n/a	0	
5 – Memorial Stones			0	+++	0	+ +				0		n/a	0	



++: Extremely positive O: Insignificant --: Extremely Negative n/a : Not computable - : Negative +: Positive

Analysis

Problems mainly with the Completeness characteristic because of its perception by users. Our definition of completeness is how complete each record is within itself. include.

Users interpret completeness as the dataset's ability to cover all of the real-life data that the dataset assumes to

Conclusion and Future Works

Interviews with willing survey participants Possible modification or replacement of metrics Possible changes to the data model

The first aim is to find more survey participants. (Please join us!)

Please consider participating in our survey or spreading the work to suitable users!

Thank you for listening.

http://tinyurl.com/OpenDataQuality