

Open Data und Datenqualität: Maschinenlesbarkeit

Tabellarische Daten mit wenigen Handgriffen für die Weiterverarbeitung fit machen

Die Idee von Open Data ist, dass Ihre Daten von anderen betrachtet, ausgewertet und auf vielfältige Art und Weise weiterverarbeitet werden können. Dabei sind das **Dateiformat** und die **Strukturierung** des Datensatzes entscheidend - diese Faktoren bestimmen, ob ein Datensatz maschinenlesbar ist.

i Ein Datensatz gilt als maschinenlesbar, wenn er von einem Computer bzw. einer Software verarbeitet werden kann.

	.pdf	.xlsx	.csv
Offenes Dateiformat	✗	✗	✓
Ideal für Menschen lesbar	✓	✓	✗
Für Programmierung/ Software geeignet	✗	u.U.	✓

Daten in Excel-Tabellen (.xlsx) sind nur unter Umständen maschinenlesbar. Wenn Sie die folgende Empfehlungen beachten, erleichtern Sie die Nachnutzung Ihrer Daten.

6 Empfehlungen, um die Datenqualität zu erhöhen

- Beschreibungen in die Metadaten**
Beschreibungen wie Titel, Autor:in oder Datum sollten nicht in der Tabelle selbst sondern in den Metadaten stehen.
- Unverarbeitete Daten**
Lassen Sie den Datensatz möglichst originär, entfernen Sie z.B. Grafiken und Berechnungen und vermeiden Sie Umbrüche.
- Eindeutige Zuordnungen der Datenfelder**
Lösen Sie verbundene Zellen auf und geben Sie jedem Datenfeld einen Wert. Idealerweise sind die Zellen und Spaltenbeschriftung einheitlich gewählt.
- Eine Tabelle pro Blatt**
Entfernen Sie Leerspalten und doppelte Spaltenbezeichnungen. Es darf nur eine Tabelle pro Blatt geben.
- Fehlende Werte eintragen**
Nutzen Sie den in der Informationsverarbeitung üblichen Wert NA (not available) um Datenfelder ohne Informationen auszuweisen.
- Standardisierte Datums- und Zeitangaben**
Verwenden Sie standardisierte Datums- und Zeitangaben wie JJJJ-MM-TT und HH:MM:SS.

In zwei Schritten von der Excel- zur .csv-Datei

Eine Excel-Datei, die die oben genannten Anforderungen erfüllt, kann ganz einfach in das offene Dateiformat .csv überführt werden. Speichern Sie dazu einfach innerhalb des Programms Ihre Excel-Datei als .csv-Datei. Beachten Sie die UTF-8-Codierung.

Excel-Datei öffnen ► **Speichern unter** ► Wählen Sie .csv und UTF-8 als Standardcodierung aus

Für noch mehr Informationen empfehlen wir den Fraunhofer Fokus Leitfaden für hochwertige Daten: www.nqdm-projekt.de