

**Catalogo dei dati e dei metadati in
possesso delle amministrazioni, che si
intendono rilasciati come dati di tipo
aperto**

Formati aperti e metadati per il riutilizzo e la diffusione dei dati pubblici

Il presente allegato illustra brevemente la definizione di formato aperto ed elenca i formati aperti più comuni ed i metadati per la pubblicazione dei dati aperti. Le informazioni contenute in questo allegato saranno aggiornate qualora la pubblicazione di particolari dati di settore e innovazioni legislative e tecnologiche lo richiedano.

Formati aperti

Come indicato nel documento delle Linee Guida, i dati di tipo aperto devono essere “accessibili attraverso le tecnologie dell'informazione e della comunicazione, ivi comprese le reti telematiche pubbliche e private, in formati aperti”, cioè resi pubblici, documentati esaustivamente e senza vincoli all'implementazione; “sono adatti all'utilizzo automatico da parte di programmi per elaboratori” (machine readable) e sono “provvisi dei relativi metadati”. L'accessibilità da parte di altri programmi di elaborazione automatica è ottenuta attraverso l'uso di formati aperti tali da permettere di massimizzare le loro possibilità di riutilizzo e interoperabilità. Inoltre il processo di creazione e pubblicazione dei dati è molto più semplice quando il formato di dato aperto viene adottato direttamente alla sorgente invece della riconduzione successiva con delle attività di trasformazione e modifica.

A titolo di esempio, la tabella 1 fornisce un primo elenco dei formati aperti maggiormente diffusi nell'ambito dell'Azienda.

Tabella 1. Formati aperti adatti

Nome (Acronimo) Descrizione	Tipo di Dato	Estensione del file
Comma Separated Value (CSV) <i>Formato testuale per l'interscambio di tabelle, le cui righe corrispondono a record e i cui valori delle singole colonne sono separati da una virgola (o punto e virgola)</i>	Dato tabellare	csv
Open Document Format per dati tabellari (ODS) <i>Formato per l'archiviazione e lo scambio di fogli di calcolo</i>	Dato tabellare	ods
Extensible Markup Language (XML) <i>E' un formato di markup, ovvero basato su un meccanismo che consente di definire e controllare il significato degli elementi contenuti in un documento o in un testo attraverso delle etichette (markup)</i>	Dato strutturato	xml

Metadati

Nello specifico ambito dei dati aperti, i metadati sono intesi come i "dati associati ai dati" tali da informare gli utilizzatori del significato del dato stesso. La definizione stessa di metadato evidenzia l'ampiezza dei possibili metadati associabili ad un insieme di dati (*dataset*), ed è per questo che è opportuno identificare un insieme minimo di informazioni di base che possano essere utilizzate ai fini della ricerca, dell'accesso e dell'utilizzo del dato pubblicato. Diventa altresì essenziale che venga posta la massima cura nel fornire tali informazioni all'utilizzatore del dato.

Alcuni dati di settore includono anche il metadato già codificato in formato standard, ma nella maggior parte dei casi è necessario definire una documentazione di minima non essendo disponibili standard precisi in tale senso. Per tale motivo è qui di seguito definito un set minimo di metadati considerati indispensabili per la pubblicazione dei dati aperti.

	consulta il dataset (es. "Descrizione' • Nome breve. E' il nome dell'attributo utilizzato nel dataset, normalmente un'abbreviazione del nome completo (es."Desc"* • Descrizione. Illustra il significato del campo, eventuali relazioni con altri dataset, come il campo e stato generato, etc. (es."Codice ISTAT del comune amministrativo' • Tipologia dell'attributo (testo, numero, data, etc.)* • Formato dell'attributo (es. unità di misura o struttura della data) • Valori ammessi: rappresenta il dominio del attributo e ne contiene, ove necessario, la decodifica (es. "E codice assume valori che vanno da 1 a 223, numero dei comuni totali della Provincia Autonoma di Trento'
Copertura Temporale (Data di inizio)	Data di inizio validità
Copertura Temporale (Data di fine)	Data di fine validità
Data di pubblicazione	Data in cui il dataset è reso disponibile on-line
Data di aggiornamento	Data di ultima modifica del dataset
Licenza	Tipo di licenza applicata
Formato	Formato di pubblicazione (standard) del dataset*
Codifica Caratteri	Codifica dei caratteri utilizzata (es. "latin-1", "PC-850"**)
Dimensione	Dimensione in byte (o multipli di byte) del dataset
URI permanente	Identificatore univoco del dataset

Catalogo dei dati

Non risultano alla data di pubblicazione del presente documento, regolamenti o convenzioni che disciplinino l'accesso a dati.