

ALLEGATO TECNICO

“DATI IN FORMATO APERTO”

1. Dati, dataset, banca dati e documenti di tipo aperto

Il patrimonio informativo pubblico è composto da tutti i contenuti prodotti da una pubblica amministrazione ed è reso disponibile per la lettura e visualizzazione, in forma scritta o altra forma (immagine, audio, video).

Le definizioni utili ai fini dell'applicazione della sopracitata normativa:

- **Dato:** è l'informazione al netto della descrizione e priva di presentazione (es. un numero, una parola, una immagine);
- **Dataset:** è una serie di dati e di descrittori degli stessi, organizzati in base a determinate caratteristiche o attributi (es. un elenco di numeri, una tabella di indirizzi, un insieme di “log” ossia informazioni specifiche dello stesso tipo utilizzate da applicazioni informatiche);
- **Banca dati o database:** è un insieme organizzato di dati opportunamente strutturato per lo svolgimento di determinate attività (es. banca dati dei bandi di gara, banca dati degli incarichi);
- **Informazione:** è un elemento che consente la conoscenza di fatti (una notizia, un progetto ma anche informazioni descrittive delle categorie di dati come i metadati). Aggiunta al dato costituisce il sistema informativo;
- **Documento:** rappresenta una insieme di informazioni che consentono la conoscenza di fatti (atti, pubblicazioni editoriali, rapporti, grafici, materiali didattici, contenuti multimediali).

2. Cos'è il formato di un file

Un formato di file è la sua estensione. Indica la convenzione che viene usata per leggere, scrivere e interpretarne i contenuti. Sono i tre (o quattro) caratteri (“.exe”, “.com”, “.html”, “.doc”) che collegano il registro del sistema operativo (tipicamente windows) con l'applicazione da utilizzare per aprirli. Ad es. microsoft word per aprire un file .doc o .rtf, Google chrome per aprire un file .html.

Sebbene i file (o flussi di dati) siano composti dagli stessi elementi (i byte), ogni file può archiviare contenuti molto diversi fra loro: testi, video, audio, immagini, ecc.

Grazie al formato è possibile interpretare correttamente i byte di un file a seconda del tipo di contenuto, e indicare al programma software come interpretare questi dati.

3. Formati proprietari e aperti

La distinzione (proprietario o aperto) può riguardare sia i sistemi operativi (es. windows, linux), sia i dati e sia le applicazioni che vengono utilizzate per produrli. A noi interessano in particolare le differenze che riguardano i documenti.

Il formato proprietario o aperto per i sistemi operativi e per le applicazioni riguarda la diffusione pubblica o meno dei codici sorgenti (algoritmi) che sono stati scritti per compilare il linguaggio macchina necessario.

Quando sono in formato aperto sono solitamente gratuiti e vengono diffusi sia i codici sorgenti che i file binari compilati. Il loro utilizzo è libero o soggetto a licenza [creative common](#).

Quando sono in formato chiuso non sono gratuiti (salvo un periodo di prova) e vengono diffusi soltanto i file binari, che sono tutelati da copyright.

I formati aperti incoraggiano la concorrenza tra sviluppatori di funzionalità software invece di garantire il controllo di un solo produttore su tutti i contenuti realizzati dagli utenti attraverso il formato proprietario.

Per quanto riguarda i dati ed i documenti (che sono quelli che ci interessano) la definizione formato aperto deve rispondere ad una serie di requisiti standard e di linee guida che devono essere rigorosamente rispettate per garantire la leggibilità, il reperimento ed il riuso dei documenti stessi.

4. L'uso dei formati tecnici aperti negli Open Data

La libera circolazione dei dati sul web si basa sulla accessibilità totale, e può essere assicurata con l'adozione di un formato tecnico aperto in grado di eliminare i problemi legati ai diritti legali e alle specifiche tecniche del software.

Il formato aperto favorisce quindi la circolazione dei dati e ne garantisce un uso più completo, dalla semplice consultazione, all'aggiornamento, al mash-up (l'incrocio con altri dati).

Per la pubblica amministrazione italiana il documento di riferimento su questi elementi è rappresentato dalle *“Linee guida per la valorizzazione del patrimonio informativo pubblico elaborate dall'AGID”*¹.

5. Formati per la pubblicazione di documenti aperti

Per quanto riguarda i documenti come delibere e atti amministrativi da pubblicare on line in Italia i formati aperti di riferimento sono²:

- l'ODT per documenti testuali,
- l'ODS per fogli di calcolo
- l'ODP per documenti di presentazione³.

Il formato più utilizzato per la rappresentazione di documenti contenenti testo e immagini che sia indipendente dalla piattaforma di lettura è il formato PDF creato da Adobe ed oggi standardizzato dall'ISO.

¹ http://www.agid.gov.it/sites/default/files/linee_guida/patrimoniopubblicol2014_v0.7finale.pdf

² L'Agenzia per l'Italia Digitale istruisce e aggiorna, con periodicità almeno annuale, un repertorio dei formati aperti utilizzabili nelle pubbliche amministrazioni, secondo quanto disposto dall'art. 68, c. 4, d.lgs. n. 82/2005.

³ Questi tre formati sono basati sullo standard XML e sono stati adottati come formato principale in alcune suite per l'automazione d'ufficio come OpenOffice.org e LibreOffice, supportata anche da altre suite come Microsoft Office, Google Drive e IBM Lotus. Nel caso in cui c'è l'esigenza di pubblicare documenti contenenti dati solo in forma tabellare, come elenchi (come ad esempio nel caso dei tassi di assenza del personale delle pubbliche amministrazioni), senza elementi aggiuntivi (e.g., grafici, formule) l'AgID consiglia di utilizzare il formato CSV. Negli altri casi di dati in forma tabellare con elementi aggiuntivi raccomanda l'uso di ODS.

Ai fini delle attestazioni OIV sull'assolvimento degli obblighi di pubblicazione, inoltre, l'A.N.AC., nella delibera n. 43/2016, allegato 4, ha chiarito che il formato di pubblicazione è **aperto** se è in formato pdf elaborabile, ods, csv, ed è **elaborabile** se è in formato xls, html.

6. Dati di tipo aperto

Il [Dpcm 3 dicembre 2013](#) detta le regole tecniche per la conservazione dei dati.

L'[art 68](#) del codice dell'amministrazione digitale definisce un dato aperto un dato pubblico:

- disponibile in forma disaggregata;
- disponibile secondo i termini di una licenza che ne consenta il riutilizzo anche per finalità commerciali;
- accessibile attraverso le tecnologie dell'informazione e della comunicazione in formato aperto (i.e., formato pubblico, documentato e neutro rispetto alle tecnologie);
- disponibile gratuitamente oppure ai costi marginali sostenuti per la sua riproduzione e divulgazione.

Si definisce, pertanto, dato aperto “un dato che può essere liberamente acceduto, utilizzato, modificato e condiviso da chiunque e per qualunque scopo, soggetto al massimo ai requisiti di provenienza (richiesta di attribuzione) e apertura (condivisione allo stesso modo)”.

Il dato di tipo aperto è quindi un dato della pubblica amministrazione, conoscibile (pubblico), a cui è associata una licenza che ne consente il libero utilizzo (disponibile) e che abbia le caratteristiche di accessibilità e gratuità come sopra definite.

Fonti:

“Formati aperti”, realizzato da Formez Pa nell'Ambito 2 Linea 2 del Programma Operativo di Assistenza Tecnica 2012-2015 (POAT), Ob.II.4 PON GAT (FESR) 2007-2013 a titolarità del Ministero dell'Economia e delle Finanze.

“Linee guida per la valorizzazione del patrimonio informativo pubblico elaborate dall'AGID”.

Allegato n. 2 delibera n. 50/2013 A.N.AC.

Allegato n. 4 delibera n. 43/2016 A.N.AC.