

# Allegato 8

## Formati raccomandati per la produzione di documenti informatici

### 1. Formati raccomandati per documenti impaginati

Il formato **PDF** è raccomandato per i documenti impaginati destinati alla spedizione e alla conservazione.

Il formato raccomandato per la produzione di documenti informatici in senso stretto (quindi, tra le altre cose, non più modificabili) il formato raccomandato è il **PDF/A-1** per via della maggiore “superficie di interoperabilità”; il PDF in generale è comunque il formato più raccomandato per i documenti impaginati, inclusi quelli che devono rimanere parzialmente compilabili o commentabili (come i moduli delle procedure amministrative).

Per quanto riguarda i formati di documenti impaginati che possono essere revisionati e modificati, o dai quali si possono derivare altri documenti, si raccomanda siano prodotti in formato **OpenDocument (.odt)**, interoperabilmente utilizzabile dai principali applicativi di videoscrittura e, laddove non sia possibile, nel formato OOXML (.docx), ma con profilo Strict. Nel caso di documenti semilavorati a carattere temporaneo e non definitivo è consigliabile anche l'utilizzo di formati, l'inclusione di documenti (anche in formati diversi) come *allegati* di un unico file PDF, l'inclusione di essenze multimediali e modelli bi- o tri-dimensionali manipolabili in tempo reale all'interno del documento, lo rendono particolarmente versatile in molteplici occasioni. puramente “virtuali” quali quelli delle suite collaborative di fornitori di servizi in Cloud qualificati.

### 2. Iperestesi

Il principale tipo di sintassi è costituito dall'*Extensible Markup Language* (XML), la cui caratteristica aggiuntiva è di essere facilmente *human readable*.

1. Si raccomanda di usare per gli iperestesi i formati più aperti, interoperabili e indipendenti dall'applicativo utilizzato, come ad esempio l'XML (con i suoi dialetti) e il markdown. Nel caso specifico di documenti destinati ad uso tramite internet o intranet, la scelta ricadrebbe naturalmente sulle versioni più recenti di HTML (HTML5) e XHTML, anche se tali linguaggi da un lato mantengono una dipendenza dal formato del documento (“pagina web”), dall'altro sono largamente dipendenti –per la loro visualizzazione– da altri file che ne descrivono la rappresentazione grafica (e.g. stili XSLT/XSLT e fogli di stile CSS). Si invita dunque ad una scelta adeguata alle finalità del documento.

2. I documenti in formato XML sono adatti alla conservazione soltanto se accompagnati dal loro schema XML (XSD). Le pagine web possono essere mandate in conservazione soltanto quando completamente statiche, combinando il “contenuto vero e proprio” in formato HTML (incluso XHTML), con la parte “presentazionale” in formato CSS. Le pagine web con contenuto dinamico (ad esempio codice JavaScript lato client) non sono adatte alla conservazione a meno di non conservare l'intero contenuto JavaScript (incluse le librerie eventualmente richiamate – cfr. §2.15) che, a sua volta, non deve riferirsi esternamente ad alcun altro documento.

### 3. Dati strutturati

1. La scelta dei formati di file per conservare dati strutturati quali grandi e piccole basi di dati è soggetta non soltanto alla tipologia dei dati “a riposo”, ma anche agli aspetti “dinamici” legati alla loro generazione e riutilizzo. Sono perciò coinvolti aspetti fortemente quantitativi sui dati, quali:

- dimensione informatica delle evidenze “a riposo”,

- capacità dei flussi informatici “in transit” (banda richiesta e sue variazioni statistiche in base alla distribuzione geografica e cronologica, valutata su più scale di grandezza e indici statistici);
  - previsioni sul ciclo di vita (generazione, modifiche, trasporto, archiviazione, distruzione);
  - considerazioni in merito a conservazione e interoperabilità in generale (a livello europeo e nazionale);
  - considerazioni in merito alla protezione dei dati, con particolare riferimento a: o dati personali e privacy: cfr. Regolamento (UE) No 679/2016 (“GDPR”) del Parlamento europeo e del Consiglio e il D.Lgs. 101/2018;
- o trattamento da parte di infrastrutture critiche e fornitori di servizi essenziali: cfr. il D.Lgs. 65/2018 e la Direttiva (UE) No 1148/2016 (“NIS”) da esso recepita.
- Si raccomanda perciò alle PP.AA. di effettuare un’adeguata valutazione di interoperabilità, che tenga in considerazione anche dei sopracitati aspetti.

#### **4. Posta elettronica**

1. Si raccomanda di utilizzare il formato EML per archiviare un singolo messaggio di posta elettronica, ovvero il formato MBOX per l’archiviazione di più messaggi ovvero di un’intera casella di posta elettronica

#### **5. Fogli di calcolo e presentazioni multimediali**

Si raccomanda la produzione di fogli di calcolo e presentazioni multimediali in formati aperti e consistenti con gli applicativi “da ufficio” più diffusi sul territorio nazionale e comunitario: in particolare, si individua nei formati derivati da OOXML (profilo Strict) e da OpenDocument le alternative più valide.

Nel caso di documenti semilavorati a carattere temporaneo e non definitivo è consigliabile anche l’utilizzo di formati puramente “virtuali” quali quelli delle suite collaborative di fornitori di servizi in Cloud qualificati, purché tali formati siano interoperabili e disponibili su sistemi informativi senza vincoli o particolari requisiti tecnologici (ad esempio, totalmente fruibili attraverso browser web).

#### **6. Immagini**

- Immagini raster [immagini che sono descritte come un insieme finito di punti (*pixel*) virtuali disposti regolarmente su una griglia rettangolare]

Per i formati generici (PNG, JPEG, TIFF, GIF) sussiste in generale l’obbligo di riconoscimento (documenti in entrata). Nonostante tale obbligo in merito ai formati generali, il Comune, qualora accetti documenti informatici in formati di immagini raster generici tra quelli sopra elencati per un procedimento amministrativo o a scopo di conservazione, può decidere di limitare ulteriormente l’accettazione di tali formati ad un loro sottoinsieme, specificatamente allo scopo di ridurre o uniformare i formati dei documenti.

Per la produzione di immagini raster, invece, la scelta dell’uno piuttosto che dell’altro formato generico, ovvero di uno di tali formati invece di uno specializzato, deve essere motivata da motivazioni tecniche ovvero da una valutazione di interoperabilità. In particolare si raccomanda:

- Il formato TIFF per immagini raster –generate o riversate in tale formato– dove la rappresentazione fedele del documento sia un vincolo tecnico o giuridico e dove la capacità di complementare l’immagine con trasparenze, impostazioni di stampa o scansione, raccordo con dimensioni fisiche di rappresentazione, etc.) costituisca un valore aggiunto.
- Il formato JPEG per la produzione di immagini originali e rettangolari il cui scopo sia meramente rappresentativo e non probatorio; il livello di compressione per la produzione di

tali immagini sarà dunque scelto in maniera adeguata a non compromettere lo scopo rappresentativo del documento.

- Il formato JPEG per il riversamento di immagini raster rettangolari ove la conservazione della qualità originale dell'immagine non costituisca un impedimento giuridico o non sia un vincolo esatto; in particolar modo si preferisca tale formato per immagini di provenienza fotografica, scegliendo anche in questo caso un adeguato livello di compressione.
- Il formato PNG per immagini raster –generate o riversate in tale formato– ove sia importante il mantenimento della qualità –rispettivamente massima o originale– solo relativamente ad una rappresentazione su schermi digitali non professionali. Sono un esempio di tale esigenza la produzione di immagini o fotografie digitali per l'utilizzo su pagine web o sulle GUI di software applicativi, così come loghi e altri simboli grafici, spesso coadiuvati da effetti di trasparenza.
- Il formato GIF per immagini raster che soddisfano i requisiti del formato precedente (PNG) salvo l'appartenenza all'ambito fotografico, ma abbiano in più almeno una delle seguenti caratteristiche tecniche:
  - utilizzo di un numero complessivo di valori colorimetrici non superiore a 256;
  - assenza di trasparenze, ovvero
  - impiego di soli due livelli di trasparenza: 'trasparenza totale' e 'opacità totale';
  - piccola animazione costituita da pochi fotogrammi, riprodotti ciclicamente o una sola volta, ove ogni fotogramma soddisfi le precedenti caratteristiche.
- Formato JPEG2000 per le immagini raster di grandi dimensioni di competenza territoriale (e.g. immagini satellitari, demaniali, catastali), urbanistico e militare.
- Formato DNG per scatti provenienti da fotocamere o cineprese digitali — siano essi generati o riversati dai formati nativi in tale formato.

Tra i formati sopra individuati, **il PNG e le varianti di TIFF** senza ausilio di compressione, ovvero con algoritmi di compressione privi di perdita (*lossless*), sono **i formati più adatti alla conservazione**. Possono essere adatti alla conservazione anche il formato JPEG ovvero le varianti di TIFF e EXR che adottino algoritmi di compressione, ma solo qualora le immagini siano state nativamente generate in codesti formati (p.es. provenienti da fotocamere ovvero scanner digitali). Sono dunque esclusi dalla conservazione i riversamenti di immagini in formati che aggiungono (o cambiano) algoritmi di compressione adottati. Gli eventuali algoritmi di compressione adottati in conservazione devono tuttavia essere algoritmi aperti, previsti e pienamente descritti nelle specifiche tecniche dei formati stesso (ovvero nelle cui specifiche tecniche sono riportati i nomi degli standard relativi a codesti algoritmi di compressione).

- Immagini vettoriali

Si raccomanda la creazione di immagini vettoriali e modelli tridimensionali in formati aperti ed interoperabili: per le immagini vettoriali sullo standard SVG del W3C viene adottato in un numero sempre crescente di contesti ed è oltretutto in continuo aggiornamento con nuove funzionalità.

## 7. Caratteri tipografici

Le seguenti famiglie di caratteri tipografici –organizzati in macro-tipologie– sono considerati “standard”

- *bastoni*: sans-serif, Arial, Helvetica, Trebuchet, Verdana, Lucida Sans, Comic Sans;
- *con grazie*: serif, Times, Times New Roman, Palatino, Georgia;
- *larghezza fissa*: monospace, Courier, Courier new, Lucida Console;
- *simboli*: Symbol, Zapf Dingbats, Webdings, Wingdings.

I caratteri tipografici di cui al punto precedente sono chiamati, all'interno delle Linee guida, "font interoperabili".

Sia per l'utilizzo applicativo che per quello web si fanno le seguenti due raccomandazioni:

- a) Qualora si adoperino caratteri tipografici standard (come indicati all'inizio della sezione) i punti seguenti non costituiscono più raccomandazioni.
- b) Si includano sempre i file dei caratteri tipografici utilizzati nel documento (quando tali caratteri possono essere imbustati nel formato del documento), ovvero si formi un pacchetto di file che comprenda anche tali caratteri tipografici.
- c) Qualora i caratteri tipografici siano utilizzati per rappresentare testi scritti (come semplice collezione di simboli), scegliere solo caratteri tipografici che contengano un numero sufficiente di glifi a rappresentare *almeno* i caratteri alfanumerici, di interpunzione e diacritici del linguaggio utilizzato, della lingua italiana e di quella inglese.
- d) Ottemperare, se possibile, alla raccomandazione (a) per i caratteri alfanumerici, di interpunzione e diacritici di tutte le lingue dell'Unione Europea;
- e) Qualora dei caratteri tipografici non standard siano utilizzati per rappresentare testi scritti, mettere in atto tutte le metodiche tecniche e operative previste dai formati di file e di pacchetti utilizzati affinché, nel caso in cui i file dei caratteri tipografici non siano disponibili, l'applicativo usato per visualizzare, stampare o riprodurre il documento possa effettuare sostituzioni dei caratteri tipografici con altri caratteri standard, senza che ciò risulti in una variazione sostanziale del contenuto informativo del documento.

## 8. Audio

Per quanto riguarda la creazione di documenti audio (di entrambe le categorie elencate all'inizio della sezione) si raccomanda la scelta di formati interoperabili ed aperti: [Broadcast] Waveform File (Estensione/i .wav, .bwf, .rf64), MPEG-1, Layer 3 (Estensione/i .mp3); Free Lossless Audio Codec (Estensione .flac) Audio "Raw" (Estensioni .pcm, .raw, .sam) MusicXML™ (Estensione .musicxml)

## 9. Video

1. I file e i flussi multimediali –e in particolar modo quelli video– contengono evidenze con caratteristica peculiare rispetto a tutte le altre tipologie di file: i flussi video (compressi o meno che siano) hanno una dimensione digitale elevata e, tipicamente, richiedono anche una banda passante minima in caso vadano formati, riprodotti o addirittura elaborati in tempo reale. Per questo motivo le scelte dei formati avranno un'elevata variabilità in base alla finalità d'uso e ai vincoli tecnologici ad esse collegati.

2. Lo strumento della valutazione di interoperabilità<sup>1</sup> può venire in contro alla risoluzione *ex ante* delle problematiche di cui al punto precedente, allo scopo di individuare prima i codec da usare per far fronte a tutte le esigenze e vincoli (normativi, amministrativi e tecnologici) in merito al documento video, durante il corso di tutto il suo ciclo vita.

3. Ciò detto, i codec da preferire per creare nuove evidenze video di uso generico sono quelli aperti e standard de iure, quali:

- MPEG4 Part-10 (colloquialmente indicato con il nome di "H.264") in quanto largamente diffuso soprattutto negli apparati di riproduzione software e hardware. Tale codec è da preferirsi per contenuti già montati e ricondotti ad una risoluzione e una qualità accettabile per tutto il ciclo vita del documento.
- MPEG2 Part-2 qualora il ciclo vita del documento video e le sue finalità d'uso impongano privilegiare la semplicità computazionale dell'algoritmo di decompressione rispetto a fattori quali la qualità o la dimensione binaria dell'evidenza.

---

1 Si rimanda all'allegato 2 delle Linee Guida Agid, §3.1

- Qualunque codec senza compressione, ovvero con una compressione priva di perdite.

Per la produzione di sottotitoli e didascalie in ambito cinetelvisivo, così come per finalità di conservazione, si raccomanda l'utilizzo del formato TTML del W3C, conforme al profilo IMSC1 e senza la specifica di caratteri tipografici esterni.

## 10. Archivi compressi

Alcuni dei pacchetti di file oggetto del presente Allegato sono di fatto creati, come ultimo passaggio, impacchettando una porzione di filesystem in un unico file compresso, utilizzando formati descritti nel presente capitolo o meno. Tali documenti sono dunque rappresentati, grazie a questo stratagemma tecnico, da un unico file anziché da un pacchetto — con indubbi vantaggi sia operativi che normativi.

Alcuni di questi formati mantengono l'estensione file del formato di archiviazione sopra descritto (e.g. .zip, .tar, .7z, ...); altri utilizzano estensioni proprie (e.g. .docx/.xlsx/.pptx, .od?, .woff/.woff2, .jar, .apk, .ipa).

Si sconsiglia di utilizzare il DCP come formato per archiviazione e conservazione.

Molto diffuso è l'uso della pacchettizzazione in formato TAR seguita da compressione, ove si adottano due *naming convention* alternative: quella con 'doppia estensione' (all'estensione .tar viene concatenata quella dello specifico algoritmo di compressione) e quella ove un'unica estensione di file indica la concatenazione dei due formati. Tale dualismo si trova, ad esempio, per le compressioni GZIP (.tar.gz ovvero .tgz), BZIP2 (.tar.bz2 / .tbz), 7-Zip (.tar.7z / .t7z), LZMA (.tar.lzma / .tlz), XZ (.tar.xz e .txz), ecc

La raccomandazione circa i formati da utilizzare per la produzione di archivi compressi non può che essere dettata dalle finalità d'uso dell'archivio. Per il resto — con l'eccezione del formato RAR e del DMG di Apple — tutti i formati elencati in questa sezione sono sufficientemente aperti. Mentre la scelta su ZIP, GZIP (e loro varianti) piuttosto che 7-Zip è dettata spesso da esigenze tecniche specifiche degli algoritmi di compressione.

## 11 Documenti amministrativi

Sono qui elencati alcuni formati di file utilizzati per documenti amministrativi di utilizzo generale, quali:

- fatturazione elettronica,
- protocollo informatico,
- asserzioni elettroniche legate a schemi di identificazione elettronica e a loro utilizzi in capo ad autenticazione, autorizzazione, sottoscrizione o altro.

Per lo svolgimento di procedimenti amministrativi non contemplati in questa sezione, l'ente può utilizzare anche altri formati di file qualora, in quest'ultimo caso, i formati non siano sostituibili con formati previsti nell'Allegato e sia comunque stata fatta una valutazione di interoperabilità. **61**

### Fattura elettronica:

Nome completo

fattura elettronica FatturaPA

Estensione/i

.xml

Specializzazione di

XML

Tipo MIME

application/xml

Sviluppato da

Agenzia delle Entrate

Tipologia di standard

aperto, estendibile, *de iure*, testuale

Livello metadati

4

Derivato da

—

Revisione

1.2.1 (2018)

Riferimenti

- [www.fatturapa.gov.it](http://www.fatturapa.gov.it)
- *Specifiche tecniche operative del formato della fattura del sistema di interscamio* v1.2.1 (2018)
- *Schema del file XML Fattura PA* v1.2.1 (2018)
- Foglio di stile per visualizzare la fattura v1.2.1 (2018)
- Agenzia delle Entrate, *Allegato A* del D.M. 55/2013

Conservazione

Sì

Racc. per la lettura

Specifico; consultare la normativa in materia

Racc. per la scrittura

Specifico; consultare la normativa in materia

La fattura elettronica è creata, trasmessa, elaborata e conservata centralmente dal Sistema di Interscambio (*SdI*) dell'Agenzia delle Entrate, ovvero da altri applicativi ad esso interconnessi. Il formato della fattura elettronica è costituito da un file XML nel dialetto FatturaPA, che comprende tutte le varianti di processo della fatturazione convenzionale, suddivisi in tre macro-aree:

- dati anagrafico-fiscali delle parti (prestatore/cedente e committente/cessionario),
- metadati relativi alla fattura elettronica stessa, alle valute e al suo trasporto nel SdI,

## Segnatura di Protocollo

Nome completo

Segnatura di protocollo

Estensione/i

.xml

Specializzazione di

XML

Tipo MIME

application/xml

Sviluppato da

Agenzia per l'Italia Digitale

Tipologia di standard

aperto, estendibile, *de iure*, testuale

Livello metadati

2

Derivato da

—

Revisione

2013

Riferimenti

- Allegato 6 alle Linee guida per la formazione, gestione e conservazione del documento informatico
- Foglio di stile per visualizzare la segnatura (2018)

Conservazione

Sì

Racc. per la lettura

Specifico; si veda la normativa nelle presenti Linee guida

Racc. per la scrittura

Specifico; si veda la normativa nelle Linee guida

## Asserzione SPID

Nome completo

Asserzione SPID

Estensione/i

.xml

Specializzazione di

XML

Tipo MIME

text/xml

Sviluppato da

Agenzia per l'Italia Digitale

Tipologia di standard

aperto, estendibile, *de iure*, testuale

Livello metadati

2

Derivato da

Security Assertion Markup Language, SSO profile; OpenID Connect

Revisione

1.0 (2014)

Riferimenti

- [www.spid.gov.it](http://www.spid.gov.it)
- AGID, *Regolamento recante le Regole Tecniche v1.0* (2014)
- D.P.C.M. 24 ottobre 2014

Conservazione

Sì

Racc. per la lettura

Specifico; consultare normativa in materia

Racc. per la scrittura

Riservato alla federazione SPID, costituita da gestori di identità digitale, fornitori di servizi e *attribute authority*.

7. Le asserzioni legate al [Sistema Pubblico di Identità Digitale \(SPID\)](#), sono da considerarsi un formato a se stante in quanto possono trasportate e, se adeguatamente gestite, conservare a lungo termine dati personali o sensibili di persone fisiche o giuridiche. Per tali motivi tali asserzioni sono normate da altre Regole Tecniche emanate dall'Agenzia per l'Italia Digitale..