Città di
Rivalta
di Torino

ALLEGATO 6
alla deliberazione della Giunta Comunale
n. 264 del 30/12/2021

FORMATI

DEI

DOCUMENTI

(già allegato 2 Linee Guida)

Formati di file e riversamento

Allegato 2 al documento
"Linee Guida sulla formazione, gestione e conservazione dei documenti informatici".

Sommario

1 Introduzione 3

1.1 Definizioni fondamentali 3
1.1.1 File, flussi digitali e buste-contenitori 4
1.1.2 Filesystem e metadati 5
1.1.3 Metadati e identificazione del formato 8
1. 2 Tassonomia 9
1.2.1 Tipologie di formati 9
1.2.2 Classificazione di formati 11
1.2.3 Formati generici e specifici 14

2. Tipi di file 16

2.1 Documenti impaginati 21
2.1.1 Raccomandazioni per la produzione di documenti 33
2.2 Ipertesti 34
2.2.1 Raccomandazioni per la produzione di documenti 41
2.3 Dati strutturati 42
2.3.1 Raccomandazioni per la produzione di documenti 52
2.4 Posta elettronica 53
2.4.1 Raccomandazioni per la produzione di documenti 55
2.5 Fogli di calcolo e presentazioni multimediali 55
2.5.1 Raccomandazioni per la produzione di documenti 59
2.6 Immagini raster 60
2.6.1 Raccomandazioni per la produzione di documenti 74
2.7 Immagini vettoriali e modellazione digitale 77
2.7.1 Raccomandazioni per la produzione di documenti 84
2.8 Caratteri tipografici 84
2.8.1 Raccomandazioni per la produzione di documenti 86
2.9 Audio e musica 87
2.9.1 Raccomandazioni per la produzione di documenti 92
2.10 Video 93
2.10.1 Raccomandazioni per la produzione di documenti 102
2.11 Sottotitoli, didascalie e dialoghi 103
2.11.1 Raccomandazioni per la produzione di documenti 108
2.12 Contenitori e pacchetti di file multimediali 108
2.12.1 Raccomandazioni per la produzione di documenti 131
2.13 Archivi compressi 132
2.13.1 Raccomandazioni per la produzione di documenti 138
2.14 Documenti amministrativi 138
2.15 Applicazioni e codice sorgente 142
2.16 Applicazioni crittografiche 142

3 Raccomandazioni sui formati di file 146

3.1 Valutazione di interoperabilità 147
3.2 Indice di interoperabilità 149
3.3 Riversamento 150

1 Introduzione

1. Il presente documento fornisce indicazioni iniziali sui formati dei file con cui vengono rappresentati i documenti informatici oggetto delle presenti linee guida. I termini indicati in azzurro, alla prima occorrenza all'interno di questo testo, sono definiti nel Glossario delle presenti Linee guida.
2. I formati descritti sono stati scelti tra quelli che possono maggiormente garantire il principio dell'interoperabilità tra i sistemi di gestione documentale e conservazione e in base alla normativa vigente riguardante specifiche tipologie di documenti. Va tuttavia segnalato che non tutti i formati di file nel presente documento sono leggibili da qualsivoglia elaboratore, a seconda della configurazione degli applicativi installati. Questo perché, nel caso di finalità specifiche e settoriali (come avviene ad esempio per i file multimediali), alcuni formati di file sono utilizzabili solo dopo l'installazione di software applicativi specifici per l'attuazione delle suddette finalità.
3. È bene precisare che, rispettando il principio di interoperabilità e cercando di mitigare il rischio di "obsolescenza tecnologica", i formati consigliati tra quelli elencati elencati –inclusi quelli per finalità specifiche, cfr. §1.2.3– sono quanto più possibile "aperti", liberamente utilizzabili e non coperti da brevetto. Sono inoltre reperibili online diversi software applicativi open-source in grado di leggere tali file. Tra i formati elencati
4. Tra i formati elencati nel presente Allegato, vi sono anche quelli non consigliati per finalità di interoperabilità, archiviazione o conservazione; essi sono presenti nell'elenco perché formati già ampiamente diffusi nella pubblica amministrazione e quindi non ignorabili per quanto riguarda il loro trattamento e il riversamento da questi formati verso formati più interoperabili.
5. Il presente Allegato, per la natura stessa dell'argomento trattato, viene periodicamente aggiornato sulla base dell'evoluzione tecnologica e dell'obsolescenza dei formati e potrà essere pubblicato online sotto forma di Avvisi, ovvero di un registro dei formati sul sito istituzionale dell'Agenzia per l'Italia Digitale[1].

1.1 Definizioni fondamentali

Si faccia riferimento al Glossario delle presenti linee guida per la definizione dei termini non ulteriormente introdotti in questa sezione.

[1] Qui di seguito indicata anche, per brevità, come Agenzia, ovvero come AGID.

1.1.1 File, flussi digitali e buste-contenitori

1. Dal punto di vista tecnologico un documento informatico è rappresentato da un file, ovvero da un flusso binario (*stream*); in linea di principio un flusso binario di dimensione finita può essere contenuto in un file. Il parametro progettuale più importante associato a un file è la sua dimensione (espressa in byte o suoi multipli). Per un flusso binario, che invece può non avere una dimensione predeterminata, si parla invece del suo *data-rate* (ovvero *bit-rate*, quando espresso in bit o suoi multipli), cioè la media temporale dei bit contenuti dal flusso nell'arco di un secondo.
2. La capacità di poter produrre, elaborare o trasmettere flussi entro un data-rate massimo attraverso un canale di comunicazione digitale costituisce la banda dell'elaboratore ovvero del canale — in inglese *bandwidth*. In questo capitolo ci interessiamo prevalentemente ai documenti informatici rappresentati mediante file, mentre sarà presa in considerazione la rappresentazione mediante flussi binari nel caso di alcuni file multimediali (cfr. §§2.9–2.12).
3. In alcuni casi il documento informatico è rappresentato da un insieme di file distinti, organizzati in un pacchetto di file, in inglese (*file*) *package*.
4. L'algoritmo che permette di rappresentare un documento informatico mediante un'evidenza quale un file tramite un'operazione di codifica, o *encoding*, definisce dunque il formato del file; l'operazione inversa, per estrapolare dai dati binari di un file –codificato in un dato formato– nel contenuto informativo del documento, è chiamata decodifica (*decoding*). Formati diversi necessitano di codificatori e decodificatori specifici, che in una parola (soprattutto quando entrambi, per un formato specifico, sono implementati da una singola componente applicativa) sono abbreviati in codec.
5. Esistono una moltitudine di formati di file per rappresentare i documenti informatici ma, a seconda del contenuto del documento e delle esigenze specifiche di gestione e conservazione dello stesso, alcuni formati sono più adatti di altri. Alcuni formati possono essere utilizzati per codificare documenti di una sola tipologia (ad es. formati di file per immagini, generalmente, non possono codificare documenti audio); altri formati, invece, vengono usati per una o più finalità di codifica tra le seguenti:

- codificare documenti di tipologie diverse (ad es. sia testi, che immagini, che audio);
- codificare più documenti insieme nel medesimo file (ad es. scopo principale dei formati di archiviazione);
- codificare documenti di una medesima tipologia (ad es. video) usando però algoritmi di codifica diversa.

I formati di file che assolvono ad una o più delle suddette funzioni sono chiamati (formati) busta, o (formati) contenitori — in inglese, rispettivamente, (*file*) *wrappers* o *containers*. Le due tipologie di documento si utilizzano prevalentemente formati contenitori sono i documenti che richiedono funzioni crittografiche avanzate (cfr. §2.16) e i file multimediali (immagini, suono, video, cfr. §§2.6–2.12).
6. Alcuni formati contenitori, infine, imbustano al loro interno, in un unico file-busta, pacchetti di più file precostituiti secondo un determinato formato di pacchetto di file. È questo il caso, ad esempio, dei formati OpenDocument e Microsoft® OOXML (cfr. §2.1 e §2.5), ovvero delle immagini virtuali di filesystem (cfr. §2.13).

1.1.2 Filesystem e metadati

1. I file sono solitamente archiviati una base di dati chiamata filesystem, ove i file con maggiore parentela fra loro (altrimenti detta *locality of reference*, ovvero "affinità per referenza") sono collocati nel medesimo nodo dell'albero: la cartella (ovvero *folder* o *directory* in inglese).
2. All'interno di un filesystem, ai file possono essere associate altre informazioni che ne completano l'esistenza all'interno dello stesso, anche se tali informazioni non fanno parte del contenuto binario del file propriamente detto; tali informazioni sono chiamate per questo motivo metadati (in inglese *metadata*) "esterni" del file.
3. Esistono molteplici formati di filesystem, che variano a seconda delle tecnologie di stoccaggio, di specifiche finalità. Alcuni di questi formati sono open-source; altri sono codificati in standard; altri ancora sono protetti da brevetti e/o copyright.
4. I metadati esterni rappresentabili in un dato filesystem possono differire anch'essi –per qualità, numero, sintassi e funzionalità– a seconda del formato di filesystem, ma di solito comprendono almeno:

- il nome del file, cioè una stringa di caratteri (di lunghezza variabile entro un limite massimo finito) che identifica univocamente il file all'interno della medesima cartella. Sussistono limitazioni differenti circa i caratteri ammessi nel nome e la sua lunghezza massima, a seconda dei diversi formati di filesystem considerati;
- la posizione virtuale del file all'interno del filesystem, chiamato "percorso del file" (path) — "locale" in quanto relativo al filesystem che lo ospita;
- la dimensione del file sopra definita, espressa da un numero intero di byte (o suoi multipli);

- la data e l'ora relativa all'ultimo istante in cui il sistema informatico che gestisce il filesystem ha rilevato una modifica del file — chiamata "data di modifica" del file.

5. La concatenazione ordinata del percorso di un file e del suo nome prende il nome di percorso completo del file (*pathname*). Invece la parte, opzionale, del nome del file costituita, scorrendo i caratteri del nome da sinistra a destra, a partire dall'ultima occorrenza del carattere punto '.' in poi (codice ASCII $2E_{16}$ in esadecimale) è chiamata –quando esiste– "estensione" del file.
6. A titolo esemplificativo, fanno parte dei metadati esterni di un file anche i seguenti:

- la data e l'ora creazione del file (che, a seconda della tipologia di filesystem può, con diversi gradi di ambiguità, coincidere con il momento di prima comparsa del file sul filesystem specifico, ovvero il momento della creazione del file sul suo filesystem di origine, o altro);
- la data e l'ora relativa al più recente accesso in lettura sul file avvenuto nel filesystem specifico;
- un identificativo più o meno univoco dell'utente informatico che è il proprietario virtuale del file (rispetto agli altri utenti virtuali del sistema) — chiamato l'*owner* del file;
- una serie di attributi che istruiscono i sistemi informativi che gestiscono l'intero filesystem di appartenenza circa la possibilità di autorizzare determinate operazioni sullo specifico file, a seconda sia dell'operazione da compiere che dell'utenza informatica che presenta tale richiesta; per ogni file, tali metadati costituiscono o sono una parte del cosiddetto insieme dei suoi permessi (*permission*s), ovvero una vera e propria lista dei controlli d'accesso (*ACL*);
- un'etichetta che stabilisce il tipo di formato file (o container) impiegato per la codifica del documento; una cui codifica universalmente riconosciuta *MIME type* – cfr. RFC-2046 e RFC-3023).

7. È importante esplicitare che le "date" sopra descritte, pur non assolvono a requisiti di integrità, precisione e immutabilità nel tempo solo per il fatto di rappresentare una data e un'ora in un filesystem, non producendo dunque, a priori, la stessa validità giuridica di marcature temporali elettroniche qualificate ai sensi del Regolamento (UE) Nº 910/2014.
8. Come anticipato in §1.1.1, il documento informatico può essere rappresentato da un insieme di file distinti, organizzati in un pacchetto di file –in inglese (*file*) *package*– ove l'affinità per referenza tra di essi è realizzata, a seconda del formato del pacchetto, mediante *una o più* delle seguenti tecniche:

- Parentela stretta dei file all'interno del filesystem realizzata definendo un sottoalbero riservano a contenente l'intero pacchetto. I file del pacchetto sono tutti contenuti nella cartella-radice del pacchetto ovvero in sue sottocartelle e, tipicamente, è escluso da tale sottoalbero qualunque file non appartenente al pacchetto. Esempi di formati che utilizzano questa tecnica sono le firme elettroniche avanzate nel formato ASiC (§2.16), il formato di master interoperabile (IMF), i pacchetti per il cinema digitale (DCP), e i pacchetti XDCAM (per tutti, §2.12).
- Sintassi dei nomi –rigida o parziale– per file ed eventuali sottocartelle costituenti il pacchetto. Formati che utilizzano esclusivamente questa tecnica sono le impronte crittografiche *detached* (§2.16) e i pacchetti video organizzati in sequenze di fotogrammi (quali ad esempio i master per la distribuzione del cinema digitale, DCDM, §2.12).
- Presenza di un "indice di pacchetto", generalmente rappresentato da un file che ha sia un nome che una posizione controllate e che contiene i *pathname* degli altri file costituenti il pacchetto. Tale indice assolve spesso a scopi aggiuntivi, come ad esempio consolidare in un unico punto i metadati esterni del pacchetto o dei singoli file (§1.1.3), che altrimenti potrebbero essere accidentalmente alterati (o persi) spostando il pacchetto da un filesystem ad un altro. In questi casi il file indice prende anche il nome di file-manifesto, ovvero *sidecar file*. Formati che utilizzano esclusivamente questo metodo sono i pacchetti di siti web (indice con nome *non* obbligatorio index.html, §2.2), alcuni tipi di firme elettroniche avanzate ASiC (tramite la cartella-indice META-INF, §2.16), i già menzionati pacchetti IMF e DCP (che utilizzano più di un file con funzioni di indice, §2.12).
- Riferimento mediante identificativi unici (UID) assegnati a ciascun file del pacchetto, quali ad esempio le loro impronte crittografiche. Esempi di formati che utilizzano questo metodo sono le firme elettroniche avanzate CAdES e XAdES *detached* (§2.16), le marche temporali *detached* e i già menzionati pacchetti IMF e DCP.
- Consolidamento di tutti e soli i file del pacchetto in un file-archivio (§2.13). Esempi che utilizzano questo metodo sono i documenti nei formati OpenDocument e Microsoft® OOXML (cfr. §2.1 e §2.5), gli applet Java e i pacchetti applicativi per dispositivi mobili con sistemi operativi Android e iOS® (§2.15), le firme elettroniche ASiC (§2.16) e, in senso lato, i file PDF (versione 1.7 e successive) quando imbustano altri documenti al loro interno sotto forma di "allegati PDF" (§2.1).

9. La tipologia di regole sintattiche che stabilisce come implementare i primi due metodi sopra elencati costituisce la *naming convention* del formato di pacchetto. La sua efficacia è ridotta quando non coadiuvata da ulteriori controlli di integrità del

pacchetto, in quanto l'affinità per referenza è generalmente difficile da far rispettare tecnicamente (a meno di usare uno stretto controllo dei permessi di "sola-lettura", ovvero archiviare su dispositivi logicamente immodificabili). Per questo motivo la *naming convention* si affianca spesso ad altri metodi quali quelli ai punti dal 3 al 5 del sopracitato elenco.

1.1.3 Metadati e identificazione del formato

1. Abbiamo già parlato in §1.1.2 dei metadati esterni, che servono a descrivere meglio un file ma sono fortemente dipendenti dal filesystem ove il file è archiviato in un dato momento. Inoltre, tali metadati possono essere soggetti a modifiche che non pregiudicano l'integrità del file stesso.
2. Più importanti ancora sono perciò i cosiddetti metadati "interni" di un file, cioè informazioni descrittive del file che sono codificate nel suo formato stesso. A seconda del formato impiegato, la presenza di questi metadati interni può essere obbligatoria o facoltativa. Facendo parte del contenuto binario di un file, la modifica di tali metadati compromette l'integrità del documento informatico.
3. Il riconoscimento (in modalità automatica o manuale) del formato di file impiegato per rappresentare un documento informatico può avvenire attraverso alcune modalità, tra cui le più diffuse sono tramite metadati interni ovvero esterni:

a) L'estensione nel nome del file, anche se tale associazione:
   - non è resiliente — è solitamente banale rinominare un file cambiandone l'estensione [ovvero crearne direttamente il nome] con un'estensione non corrispondente al formato utilizzato (es. un file di testo semplice con codifica ASCII denominato con estensione .doc, che invece è prerogativa dei documenti di Microsoft® *Word*);
   - non è biunivoca — una medesima estensione può essere usata nel nome di file codificati in formati diversi (es. l'estensione .log, usata per rappresentare file di registro codificati in maniera diversa), ovvero un dato formato viene associato a file con una molteplicità di estensionI (es. i certificati elettronici in formato X.509 con codifica DER, rappresentati con diverse estensioni tra cui .crt ovvero .cer).

b) La "tipologia MIME" (*MIME type*) del formato di file, anche se tale associazione gode di svantaggi simili all'estensione, in quanto:
   - soprattutto se espressa come metadato esterno (si veda §1.1.2), non è resiliente a variazioni o rimozioni del medesimo in maniera non controllata;

- sia nel caso in cui sia espressa come metadato interno che come metadato esterno del file, potrebbe descrivere il formato in modo comunque ambiguo.

c) La presenza di metadati interni al file, espressi in "campi" che si trovano in posizioni specifiche (prefissate ovvero ricalcolabili) all'interno del file[2]. La lettura di tali campi permette di dedurre il formato del file in maniera più diretta e affidabile. Molti formati impiegano, specificatamente a questo scopo, una stringa prefissata di pochi caratteri (generalmente dai 2 ai 6), posta all'inizio del file, chiamata *magic number* e che ne identifica univocamente[3] il tipo di file.

1. 2 Tassonomia

1.2.1 Tipologie di formati

1. L'evolversi delle tecnologie e la crescente disponibilità e complessità dell'informazione digitale ha indotto la necessità di gestire sempre maggiori forme di informazione digitale (testo, immagini, filmati, ecc.) e di disporre di funzionalità specializzate per renderne più facile la creazione e la modifica.
2. Questo fenomeno porta all'aumento del numero di formati disponibili e dei corrispondenti programmi necessari per codificarli, decodificarli e gestirli in ogni modo.
3. Segue una sommaria e non esaustiva catalogazione dei più diffusi formati di file e pacchetti, secondo il loro specifico utilizzo ("tipologia"). A fianco di ogni tipologia di formati sono indicati i formati pertinenti oggetto del presente Allegato; qualora l'estensione di file associata al formato sia diversa –a meno di maiuscole/minuscole– dall'eventuale acronimo del nome del formato stesso, essa sarà indicata affiancata al nome tra parentesi (e.g., il formato PDF non avrà un'estensione indicata tra parentesi in quanto la sua estensione predefinita è già .pdf).

- Documenti impaginati (§2.1) — PDF, Microsoft® OOXML (.docx) e *Word* (.doc), OpenDocument Text (.odt), Rich-Text Format (.rtf), EPUB, PostScript™ (.ps), Adobe® *InDesign*® Markup Language (.idml);

[2] La parte iniziale ovvero quella terminale di un file contengono spesso gran parte dei campi utili a contenere i metadati interni (e quindi anche a identificare il formato) del file; quando presenti, queste parti sono chiamate, rispettivamente, *header* (impropriamente tradotto come "intestazione") e *footer* del file.
[3] Il magic number può anche identificare l'allineamento delle word, che le architetture dei microprocessori e i sistemi operativi implementano diversamente per varie ragioni. Alcuni magic number notevoli sono indicati, per i rispettivi formati, nel §2.

- Ipertesti (§2.2) — XML, dialetti e schemi XML (.xsd, .xsl), HTML (.html, .htm), fogli di stile per XML/HTML (.xsl, .xslt, .css), Markdown (.md);
- Dati strutturati (§2.3) — SQL, CSV, Microsoft® OOXML (.accdb) e *Access* (.mdb), OpenDocument Database (.odb), JSON, Linked OpenData (.json-ld), JWT[4];
- Posta elettronica (§2.4) — .eml, .mbox;
- Fogli di calcolo (§2.5) — Microsoft® OOXML (.xlsx) e *Excel* (.xls), OpenDocument Spreadsheet (.ods);
- Presentazioni multimediali (§2.5) — Microsoft® OOXML (.pptx) e *PowerPoint* (.ppt), OpenDocument Presentation (.odp);
- Immagini raster (§2.6) — JPEG (.jpg, .jpeg), TIFF (.tif, .tiff), PNG, GIF, OpenEXR (.exr), JPEG2000 (.jp2k, .jp2c, .jp2), DICOM, Adobe® DNG, Adobe® *Photoshop*® (.psd), DPX, ARRIRAW (.ari);
- Immagini vettoriali e modellazione digitale (§2.7) — SVG, Adobe® *Illustrator*® (.ai), Encapsulated PostScript™ (.eps);
- Modelli digitali (§2.7) — StereoLithography (.stl); Autodesk® DWG™, DXF™, DWF™, FBX™.
- Caratteri tipografici (§2.8) — OpenType (.otf), TrueType (.ttf), Web Open Font (.woff, .woff2);
- Suono (§2.9) — Waveform RIFF / Broadcast Wave (.wav, .bwf), MP3, audio RAW (.pcm, .raw, .snd), AIFF (.aiff, .aifc, .aif), FLAC, MusicXML™ (.music.xml), MIDI (.mid); molteplici codec audio;
- Video (§2.10) — formati video delle famiglie MPEG2 e MPEG4; molteplici codec video;
- Sottotitoli (§2.11) — TTML/IMSC/EBU-TT (.ttml, .dfxp, .xml), EBU STL;
- Contenitori multimediali (§2.12) — MP4, MXF, MPEG2 Transport/Program Stream (.vob, .ts, .ps), AVI RIFF (.avi), Matroska (.mkv), QuickTime (.mov, .qt), WebM;
- Pacchetti multimediali (§2.12) — pacchetto di master interoperabile (IMF, IMP); pacchetto per il cinema digitale (DCP); master per la distribuzione cinematografica (DCDM); pacchetti Digital Intermediate basati su sequenze di fotogrammi (.exr/.dpx; .wav), ACES metadata file (.amf); pacchetto XDCAM;
- Archivi compressi (§2.13) — TAR, ZIP, GZIP, 7-Zip (.7z), RAR, TAR compresso (.tgz, .t7z, …), ISO9660 (.iso), VMware® Disk (.vmdk), Apple Disk Image (.dmg);
- Documenti amministrativi (§2.14) — fattura elettronica, fascicolo sanitario elettronico, *response* SAML SPID, segnatura di protocollo;

[4] Il *Java Web Token* (*JWT*) è in realtà un formato di flusso digitale, che può essere banalmente contenuto in un file, cfr. §2.3.

- Applicazioni e codice sorgente (§2.15) — eseguibili Microsoft® (.exe, .com), *applet* Java (.jar); pacchetti applicativi *Windows*® (.msi), Android (.apk), macOS® (.pkg), iOS® (.ipa); librerie statiche (.a, .lib) e dinamiche (.so, .dll, .dylib); script interpretabili (.sh, .?sh, .bat, .cmd, .py, .perl, .js, .go, .r, …); codice sorgente in vari linguaggi di programmazione (.c, .cpp, .h, .java, .asm, …).
- Applicazioni crittografiche (§2.16) — certificati elettronici (.cer, .crt, .pem), chiavi crittografiche (.pkix, .pem), marcature temporali elettroniche (.tsr, .tsd, .tst), impronte crittografiche (.sha1, .sha2, .md5, …); per le firme e i sigilli elettronici avanzati: buste crittografiche XAdES (.xml), CAdES (.p7m, .p7s), PAdES (.pdf), contenitori ASiC (.zip); KDM (.kdm.xml).

1.2.2 Classificazione di formati

1. L'evolversi delle tecnologie e la crescente disponibilità e complessità dell'informazione digitale ha indotto la necessità di gestire sempre maggiori forme di informazione digitale (testo, immagini, filmati, ecc.) e di disporre di funzionalità specializzate per renderne più facile la creazione e la modifica.
2. Gli *standard* tecnologici vengono incontro a tali esigenze, permettendo di definite regole di codifica e decodifica di un documento informatico, affinché sia rappresentato tramite un file, un flusso digitale, ovvero un pacchetto di file (tutti oggetti del presente allegato). Gli standard tendono a stabilizzare le specifiche tecniche dei formati di file –sia nel tempo che rispetto alle tecnologie di produzione, trasmissione e archiviazione– ma la loro importanza nel cristallizzare tali specifiche in una forma precisa serve ad impedire la nascita di varianti "esotiche" o dialetti non controllati del medesimo formato che, alla lunga, ne riducano l'interoperabilità (cfr. più avanti).
3. Un esempio su tutti: la mancata standardizzazione dei codec usati nei formati-busta multimediali (§2.12), specialmente dei file video, che spesso conduce all'impossibilità di riprodurre un filmato ritenuto compatibile con il sistema informativo a causa del riconoscimento della sola estensione del file da parte di svariate applicazioni (che al massimo può servire a identificare il formato contenitore), senza invece un'adeguata verifica del possesso dei codec adatti a riprodurne le essenze.
4. Gli standard migliori sono quelli che possono essere facilmente estesi, rivisti o aggiornati nel tempo per adattarsi all'immanente obsolescenza tecnologica. Tra questi inoltre, eccellono gli standard che sono *ab initio* disegnati con il preciso scopo di evolvere a lungo termine; per questo motivo essi sono detti formati "compatibili in avanti" o anche "*future-proof*".

5. Esempi di questi standard "virtuosi", sono l'XML e il JSON; il PDF e l'OpenDocument per i documenti impaginati; il TIFF (e il DNG), il PNG e il DPX per le immagini *raster*; l'SVG e il DXF per i modelli vettoriali; il TTML per i dialoghi; l'MXF e l'MP4 come contenitori multimediale; l'IMF come pacchetto di file multimediali.
6. I formati di file sono soggetti a revisioni allo scopo di includere caratteristiche evolutive; a questo scopo gli standard corrispondenti vengono "versionati" con diverse tassonomie, cioè incrementandone progressivamente i numeri delle revisioni (possibilmente in gerarchie di non più di 2-4 componenti numeriche– ovvero utilizzando l'anno o una data più precisa di riferimento, ovvero mediante sia numeri progressivi che date).
7. Quando non sono più fatti evolvere in favore di un formato differente (e quindi vengono progressivamente abbandonati), il formato diventa quiescente ed eventualmente viene "deprecato" in favore di un altro standard.
8. Sebbene sia tecnicamente possibile continuare a produrre e riprodurre file in un formato deprecato, il progresso tecnologico, o per meglio dire l'*obsolescenza tecnologica*, condannerà tali file a non essere più definitamente leggibili, a causa della mancanza di applicativi che ne implementino la decodifica (a meno di non coinvolgere risorse economiche solitamente sproporzionate al bisogno di accesso ai dati contenuti). Quando ciò accade ed è riconosciuto in maniera manifesta, si parla di formati "obsoleti".
9. Esistono varie classificazioni per le tipologie di formati, le quali sono spesso determinanti –ancor più delle loro peculiarità tecniche– per la scelta d'uso di un formato in favore di altri, ovvero per prendere decisioni ragionate circa il riversamento da un formato verso un altro (cfr. §3.3).
10. Tali classificazioni, elencate brevemente qui sotto, sono per lo più dicotomiche (cioè del tipo presenza o assenza di una determinata caratteristica) e indipendenti fra loro. Accanto ad ogni classificazione è riportato un modificatore in grassetto, il cui significato è riferito all'indice di interoperabilità introdotto nel §3.2. Un formato più essere:

a) uno standard *de facto* (**+2**) quando questioni contingenti (anche fa loro correlate), quali l'efficienza in casi d'uso reali, l'autoregolazione dei mercati di riferimento, l'efficacia tecnica, ne hanno determinano una larghissima e non trascurabile diffusione, per lo meno in settori di riferimento. Un formato può essere invece uno standard *de iure* (**+3**), quando esistono normative che ne obblighino, o per lo meno ne raccomandino, l'uso in determinati contesti amministrativo-legali e settori di riferimento. Rappresentano esempi di tali normative le Linee guida di cui il presente Allegato è parte integrante (in quanto già Regole Tecniche) ovvero, a livello comunitario, alcune Decisioni di Esecuzione (UE). Sono standard *de iure*, inoltre, tutti i formati codificati come tali dalle organizzazioni nazionali, comunitarie e internazionali che hanno il compito di definire standard e linee guida nei settori di riferimento

dei formati stessi (come, ad esempio, ISO, ITU, UNI, CEN, SMPTE, ecc.). Infine, un formato può anche non rappresentare alcuno standard (**0**), ma tale caratteristica tendenzialmente lo esclude dall'elenco dei formati virtuali contenuti nel presente Allegato.

b) aperto (**+3**) ovvero chiuso (**0**) a seconda che esista o meno, e sia resa pubblicamente disponibile, una "specifica tecnica" del medesimo: la documentazione che descrive dettagliatamente, come minimo, la procedura di formazione e di lettura di file in quel formato e, possibilmente, l'elaborazione e i suoi possibili scenari di utilizzo, spesso descritti organicamente mediante *operational patterns* (in italiano: schemi operativi).

c) proprietario (variabile) o non proprietario (**+4**), a seconda che sia stato creato da un'organizzazione privata –che dunque ne detenga la proprietà intellettuale– ovvero quando la gestione delle sue specifiche non è controllata in tale ambito (quindi possibilmente rilasciata al pubblico dominio, o comunque gestita da un organismo di standardizzazione). In particolare, i formati proprietari possono essere liberi (**+3**) ovvero limitati; in quest'ultimo caso la limitazione potrebbe permettere soltanto l'utilizzo libero di file già codificati in tale formato ma non la produzione di nuovi file (**+2**), ovvero limitare anche la lettura dei file formattati secondo tale formato (**0**); potrebbero essere possibili anche altri tipi di vincoli (pagamento di *royalty*, sottoscrizione di contratti di riservatezza o contratti vincolanti relativamente a particolari utilizzi quali lo sfruttamento commerciale, ecc.).

d) estendibile (**+2**) o non estendibile (**0**) qualora esso sia stato concepito *ab initio* per ammettere revisioni che ne aumentino progressivamente le funzionalità. I formati *non* estendibili, quindi, possono comunque essere soggetti a revisioni, che però potrebbero, per tali formati, richiedere una re-ingegnerizzazione o un adattamento più difficoltoso rispetto a formati estendibili, probabilmente anche a scapito delle compatibilità di cui al punto precedente.

e) livello del modello per i metadati (da **0** a **+3**) che segue l'analoga classificazione emanata nelle *Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico*, emanate dall'Agenzia per l'Italia Digitale, ove al livello 1 vengono attribuiti 0 punti e così via via fino al livello 4 cui sono attribuiti 3 punti. Tale valore è indicato, per i formati descritti al §1.2.3, nelle loro tabelle riassuntive.

f) non robusto (**0**) ovvero parzialmente robusto (**+1**) a seconda che il file comprenda meccanismi per verificare l'eventuale perdita di *integrità* di un file (o pacchetto di file); completamente robusto (**+2**), invece, qualora tale meccanismo sia presente e consenta, inoltre, di leggere correttamente le parti integre del file.

g) dipendente (**0**) ovvero indipendente dal dispositivo (**+4**) a seconda che esso richieda o meno specifici componenti hardware, firmware o software per essere creato o letto.
h) i formati il cui standard prevede *by design* che un applicativo in grado di interpretare una data revisione possa anche leggere file formattati con revisioni precedenti (eventualmente entro un limite massimo) si dicono retro-compatibili. Quelli per cui gli applicativi disegnati al momento in cui una data revisione sia corrente possano leggere anche file formattati in base a revisioni successive del medesimo standard si dicono invece "compatibili in avanti"[5] (**0**).
i) testuale (**0**) se, rappresentando ogni word di un file come caratteri testuali, sia possibile estrapolarne il contenuto informativo tramite lettura manuale e non automatizzata di tali caratteri — a seguito di uno sforzo di interpretazione di entità variabile, ma comunque proporzionato alle capacità intellettive di un tecnico di settore. Si parla, in alternativa, di formato binario (*binary*) (**0**) quando il processo è generalmente possibile solo mediante interpretazione automatizzata, "bit a bit", del contenuto digitale del file da parte di un algoritmo di parsing.

1.2.3 Formati generici e specifici

1. In ottica di interoperabilità è necessario, per le PP.AA., individuare un elenco di formati di file (e, *mutatis mutandis*, di formati di contenitori, pacchetti di file, flussi digitali e codec) per i quali vi siano obblighi o raccomandazioni in merito al riconoscimento o alla produzione in tali formati. Tale elenco non può presentare un unico tipo di obblighi e raccomandazioni, in quanto le organizzazioni di categoria, avendo necessità amministrative e operative differenti, utilizzano tecnologie digitali –e di conseguenza formati di file e applicazioni che li elaborano– estremamente variegate.
2. È auspicabile che i formati raccomandati nel presente Allegato siano adottati con modalità analoghe anche dagli enti privati –quando non già obbligatori per effetto di altre leggi– allo scopo di incrementare l'interoperabilità nello scambio di documenti informatici tra settore pubblico e privato. Per questo motivo queste Linee guida individuano, dall'insieme di formati di cui al §2:

[5] A titolo di esempio, un formato potrebbe essere progettato affinché, in uno schema di versionamento del tipo "*m.n*" (ove *m* sia il numero 'maggiore' della versione e *n* il numero 'minore') vi sia retrocompatibilità totale in scrittura di almeno *k* versioni precedenti, nonché estendibilità (in lettura) per tutte le revisioni minori della stessa versione. Ad esempio, per $k=2$, questo significa che un file codificato nella versione 3.12 di un dato formato, ad esempio, potrà essere prodotto anche da applicativi che producano file codificati con le versioni 4 e 5; inoltre ad un applicativo in grado di leggere la versione 3.12 sarà richiesto di poter leggere file di qualunque versione 3 (anche se eventuali funzionalità aggiuntive introdotte in revisioni minori successive alla 3.12, pur non "onorabili" dall'applicativo, non dovranno pregiudicarne la possibilità di aprire il file).

a) una categoria generale di formati, rispetto ai quali *tutte* le PP.AA. e le organizzazioni sul territorio nazionale,
    a. sono in grado di leggere file prodotti in questi formati,
    b. seguono le indicazioni riportate nella tabella riassuntiva del formato per quanto concerne la produzione di nuovi documenti in questi formati;
b) diverse categorie specifiche o speciali, diversificate in base al settore merceologico o alla natura del documento informatico rappresentabile, verso le quali sono identificati ulteriori obblighi e raccomandazioni solo verso le PP.AA. che, in quanto utenti professionali,[6] trattano documenti di quella tipologia nell'ambito ristretto, o comunque delimitato, di quella categoria specifica. Tali categorie possono essere di riferimento anche per soggetti privati che trattano professionalmente i medesimi documenti informatici.

3. L'appartenenza di un formato di file alla categoria generale piuttosto che a una o più categorie speciali è indicata nelle tabelle riassuntive dei formati contenute nel §2, possibilmente differenziata per la lettura ovvero la scrittura (cioè la formazione) di documenti in tali formati.
4. Come indicato nelle Linee guida di cui il presente Allegato è parte integrante, qualora un formato di file sia indicato come generico, ma una sua particolare applicazione in un determinato settore specialistico fornisca anche un formato di file specifico, ovvero una variazione strutturale del medesimo formato generico (ad es. dialetti, specializzazioni, varianti, schemi operativi, profili, specifiche tecniche particolari o aggiuntive), la specializzazione del formato prevale su quella generale.
5. Ad esempio, nel caso di un documento informatico contenente informazioni fiscali relative a una prestazione, cessazione di beni o servizi, le organizzazioni non devono produrlo in un qualsiasi tipo di file XML (che è altresì un formato generico per dati strutturati, anche di codesto tipo, cfr. §2.3); ne è garantita, invece la produzione sotto forma di fattura elettronica nello specifico formato FatturaPA (cfr. §2.14), individuato dal legislatore come il dialetto XML per quell'utilizzo specifico.
6. Si precisa che, sempre in analogia con l'indice di interoperabilità introdotto nel §3.2, il caso di pacchetti e contenitori di file va sempre considerato congiuntamente con tutte le sue parti costituenti: cioè, per i pacchetti di file, il formato del pacchetto in se, insieme ai formati di tutti i file costituenti il pacchetto stesso; per i contenitori di file, il formato della busta in se, insieme ai formati di tutti i flussi e i codec di tutte le essenze imbustate nel contenitore.

[6] Il termine "utente professionale" va inteso ai sensi del Regolamento (UE) № 1807/2018 (regolamento "FFnpD") del Parlamento europeo e del Consiglio relativa a un quadro applicabile alla libera circolazione dei dati non personali nell'Unione Europea.

7. Ad esempio, si consideri un contenitore multimediale[7] i cui formati della busta (e.g. MP4, §2.12) e quello delle essenze audio e sottotitoli (e.g., rispettivamente, WAVE e EBU STL, §2.9) sono contenute nella categoria generale, mentre il codec usato nell'essenza video (e.g. XDCAM EX) non è presente in nessun elenco, trattandosi di un codec chiuso e proprietario. In effetti un documento elettronico così formato non è riproducibile integralmente, a meno di non avere a disposizione nell'applicativo d'utilizzo (riproduttore multimediale) il codec proprietario XDCAM, la cui disponibilità pubblica è, ad oggi, limitata e potrebbe ridursi sensibilmente in futuro, a causa dell'obsolescenza del codec rispetto alle tecnologie degli applicativi, della scadenza di licenze d'uso e brevetti, nonché delle sorti dell'organizzazione che li detiene. In questo esempio, dunque, il documento elettronico non può considerarsi un file codificato, nella sua interezza, in un formato interoperabile: lo sono le tracce audio, sottotitoli e il contenitore che le imbusta, ma non lo è la traccia video e, quindi, non lo è il documento informatico nella sua interezza.

2. Tipi di file

1. Si faccia riferimento al Glossario delle presenti linee guida per la definizione dei termini non ulteriormente introdotti in questa sezione.
La sezione mantiene la medesima suddivisione per tipologie di formato delineata in §1.2.1. Per ciascun formato di file, pacchetto di file, busta o codec, ne vengono ricapitolate le caratteristiche principali in una tabella sinottica, recante:

- **nome abbreviato**, **acronimo** ovvero **pseudonimo** del formato (in alto a sinistra).
- TIPOLOGIA DI FORMATO (in alto a destra) — in particolare se si tratti di un formato di file, di busta/contenitore, di pacchetto di file, di flusso, ovvero un codec per essenze multimediali.
- icona del formato (in alto a destra) — riferimento visivo, a scopo di mera indicazione o suggestione, della riproduzione grafica di un'icona specifica per

[7] Questi file sono, comunemente ma in maniera inesatta, chiamati «file video», perché l'essenza più importante contenuta al loro interno è quella video, ovvero in questo caso «file 'MP4'» in quanto si esplicita il solo formato della busta, ignorando i codec delle essenze al suo interno, perché la busta è spesso –impropriamente– il solo oggetto cui venga una visibilità tecnica, verso l'utente, da parte del sistema operativo e di archiviazione che conserva e processa il file. Tuttavia, tali contenitori possono avere più tracce video, audio, sottotitoli — talvolta persino file di altri formati in allegato. L'integrità di tali file, che ne consente un utilizzo pieno sotto ogni aspetto tecnico, amministrativo e giuridico, è legata perciò alla possibilità di aprire e decodificare non solo la busta e una traccia video, ma tutte le essenze ivi contenute. Per questo motivo è importante sottolineare la presenza di essenze di più tipologie diverse al suo interno, ciascuna potenzialmente codificata con un codec diverso, da cui l'esigenza di nominare correttamente questi file come «contenitori multimediali».

la tipologia in oggetto (così come implementata nelle GUI dei principali sistemi operativi o software applicativi, per la riproduzione di detto formato).

- nome completo del formato.
- estensione di file — se più di una, la prima è da considerarsi raccomandata per la generazione dei file, mentre le altre sono estensioni meno comunemente usate, per le quali si raccomanda la capacità di lettura in tale formato. Nel caso dei pacchetti di file, sono invece indicate le estensioni dei formati di file che compongono obbligatoriamente il pacchetto (escludendo tutte le tipologie di formati ammessi dal pacchetto ma che non devono essere obbligatoriamente presenti all'interno).
- Tipologia MIME (*per i formati di file e busta*) — per i formati di file vengono indicati i possibili tipi MIME, come definiti in RFC-6838, per l'identificazione del formato a prescindere dall'estensione; qualora presenti più di un tipo MIME, il primo è sempre da intendersi quale preferenziale (e dunque raccomandato per la creazione di file in questo formato). Nel caso di pacchetti di file, sono invece indicati i tipi MIME usati dai file che compongono il pacchetto. Una lista aggiornata di tipi MIME registrati si può trovare al seguente indirizzo internet: www.digipres.org/formats/mime-types.
- Derivato da (*opzionale*) — eventuali formati di file, buste, pacchetti o codec di cui il formato è un'evoluzione ovvero, come nel caso di XML, ne rappresenti un dialetto. Possono essere indicati qui, per contenitori e pacchetti di file multimediali, l'applicabilità a uno o più schemi operativi.
- *Magic number* (*per i formati di file e busta*) — L'eventuale codice, definito nel Glossario e in §1.1.3, viene indicato nella sua codifica ASCII così; qualora i caratteri del *magic number* contengano caratteri speciali non alfanumerici e non diacritici, essi potranno essere indicati, *in toto* o in parte, in notazione esadecimale, come ad esempio 0x1A3F (2 byte).
- Profili (*per i soli codec*) — Molti codec posseggono parametri di compressione che sono usati per "tarare" l'algoritmo a diverse esigenze. Alcuni di questi parametri possono essere impostati –in maniera più o meno rigida– su valori determinati *a priori*, andando a costituire delle pre-selezioni che prendono il nomi di "profili", "livelli" o altro (a seconda del tipo di codec). Se rilevante o vincolante a livello di codec, può essere indicata qui la compatibilità di un codec con uno più schemi operativi (anche se uno schema operativo viene tecnicamente rafforzato al livello di contenitore o di pacchetto di file). Quando utile per distinguere ulteriormente i tipi di profilo fra loro, da un punto di vista puramente grafico, possono essere usati anche alcuni colori.
- Codice FourCC (*per i soli codec*) — definito in RFC-2361 (oltre che in www.fourcc.org) e nel Glossario, è usato per indicare una stringa identificativa di al più 4 caratteri alfanumerici minuscoli, indicata così, per distinguere i

codec audio e video usati nelle essenze per alcuni tipi di formati di busta multimediale (§2.12).

- Sviluppato da — Nome dell'organizzazione di standardizzazione, azienda, dicasterio governativo di competenza, società o comunità che ne detiene la proprietà intellettuale (formati proprietari), ovvero che ne mantiene lo sviluppo (formati non proprietari).
- Tipologia di standard — viene brevemente elencato se il formato sia codificato in qualche tipo di standard, se sia o meno aperto, proprietario (con eventuali licenze d'uso), testuale/binario, retrocompatibile, estendibile, robusto e dipendente dal dispositivo. Qualora non espressamente indicato in tale casella (rispetto alla classificazione delle tipologie di formato di cui al §1.2.1), viene implicitamente assunto che il formato sia:
    a) non codificato in alcuno standard, né *de iure* né *de facto*,
    b) chiuso,
    c) proprietario con licenza d'uso vincolante sia in lettura che in scrittura,
    d) non estendibile,
    e) possegga modello per i metadati di livello 1,
    f) non robusto,
    g) *indipendente* dal dispositivo.
- Livello metadati — Il "livello del modello per i metadati", come introdotto nelle *Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico* emesse dall'Agenzia per l'Italia Digitale, che si riferisce alla classificazione dei metadati supportati nativamente dal formato. Si precisa che la classificazione di ciascun formato in un dato livello si riferisce all'opportunità di valorizzare *tutti* i metadati, obbligatori e facoltativi, come previsti dalle specifiche tecniche del formato stesso riportate nella tabella, alla sezione Riferimenti). Qualora non vengano valorizzati tutti i metadati facoltativi, la classificazione può attestarsi a un livello inferiore. Altresì, non viene considerata, ai fini di tale classificazione, la valorizzazione *custom* di eventuali metadati "riservati" a casi d'uso non riportati dalle specifiche (quali, ad esempio, campi riempiti da singoli fornitori con metadati di loro propria rilevanza e sintassi). Nel caso di pacchetti di file, la classificazione si riferisce ai metadati complessivamente presenti nell'insieme di file minimi costituenti il pacchetto.
- Revisione — viene indicata qui la versione dello standard di riferimento per le PP.AA.; la "versione di riferimento" va intesa, salvo indicazione contraria in Revisione, come la versione più recente tra quelle supportate, nel senso che:

- per le organizzazioni che *leggono* file in questo formato, è obbligatoria la leggibilità di file creati con questa versione e, con tutte le precedenti con cui tale versione è descritta come retrocompatibile (cfr. §1.2.2);[8]
- per le organizzazioni che *producono* documenti in tale formato, è obbligatoria la produzione di nuovi documenti in tale versione *ovvero* – qualora non tecnicamente praticabile– con una qualunque versione precedente rispetto alla quale la versione indicata in Revisione sia descritta come retrocompatibile (cfr. §1.2.2[9]), *ovvero* –qualora non tecnicamente praticabile neanche ciò– con una qualunque versione successiva rispetto alla quale, però, la quella indicata in Revisione sia descritta come compatibile in avanti.

- Riferimenti — sono qui elencate tutte le normative di riferimento (sotto forma di leggi, regolamenti tecnici, linee guida o standard) nonché –in mancanza o a complemento di altro– *best pratices* e indirizzi di siti web ove sia resa disponibile altra utile documentazione ufficiale;
- Conservazione — Sono date indicazioni in merito all'utilizzo del formato per la conservazione di cui alle presenti Linee guida. In alcuni casi, il formato può essere adottato per la conservazione dei documenti informatici purché siano adottate specifiche configurazioni, eventualmente indicate qui. Infine, è indicato in questo campo (con la dicitura "cfr. §2.8"), l'eventuale attenzione all'uso di caratteri tipografici non interoperabili.
- Racc. per la lettura — raccomandazioni per le PP.AA. relativamente alla capacità di leggere documenti informatici nel dato formato, oltre che eventuali obblighi normati dalle Linee guida di cui questo Allegato è parte integrante;
- Racc. per la scrittura — raccomandazioni per le PP.AA. relativamente alla capacità di formare documenti informatici nel dato formato, oltre che eventuali obblighi previsti dalle Linee guida di cui questo Allegato è parte integrante. In tale sezione sono collocate anche indicazioni in merito all'uso di tale formato per la conservazione, così come trattata nelle presenti Linee guida.

Ciascun formato può essere ulteriormente descritto, particolarmente riguardo alle raccomandazioni e agli eventuali obblighi, in una parte discorsiva successiva alla tabella sinottica.

---

[8] Se, ad esempio, è raccomandata la versione 3.0 di un dato formato di file, che è definito come una variante della versione 2.0 tale che qualunque lettore in grado di leggere la 3.0 leggerà anche la 2.0, l'obbligo in lettura si applica anche alla versione 2.0.

[9] Se, utilizzando il precedente esempio, è raccomandata la versione 3.5 di un dato formato di file ma qualunque applicazione in grado

2. Si rimanda al Regolamento (UE) Nº 679/2016 ("GDPR") del Parlamento europeo e del Consiglio in materia di trattamento dei dati personali che, rappresentati in un documento informatico, possono essere soggetti a *pseudonimìa*, come metodologia atta a proteggerli, aggiornarli, cancellarli. La scelta dei formati di file da utilizzare può discendere anche dall'implementazione di tecnologie e procedure allo scopo di ottemperare agli obblighi imposti in capo al GDPR.
3. Si rimanda al Regolamento (UE) Nº 1807/2018 (regolamento "FFnpD") del Parlamento europeo e del Consiglio relativa a un quadro applicabile alla libera circolazione dei dati non personali nell'Unione Europea, per quanto riguarda ulteriori considerazioni in merito sia alla localizzazione geografica dei documenti informatici, alla loro portabilità e alle distinzioni tra utilizzo e utenti professionali di dati non personali.
4. Si rimanda alle *Linee guida sull'accessibilità* e alle *Linee guida di design* per le considerazioni sulla rappresentazione di elementi testuali e grafici nei documenti elettronici, cui le PP.AA. si attengono per la produzione di documenti informatici. Come mero esempio, tali considerazioni possono riguardare i criteri di scelta:

- delle famiglie di caratteri tipografici (§2.8) e del loro corpo;
- dell'inclusione di un elemento grafico come immagine vettoriale (§2.6) ovvero *raster* (§2.7).

5. Il primo criterio di scelta è di particolare pertinenza con la formazione dei documenti informatici, in quanto in moltissime tipologie elencate nel §2 (non solo i documenti impaginati, bensì anche le pagine web, i fogli di calcolo, le presentazioni multimediali, le immagini vettoriali e la modellazione digitale, nonché i sottotitoli) il tipo di carattere è parte integrante del documento perciò, qualora il carattere tipografico non sia presente o non sia interpretabile dal sistema informativo che visualizza il documento, ciò può comportare una difformità o addirittura l'impossibilità di leggere lo stesso.
6. Nel caso di documenti la cui rappresentazione o utilizzo (si pensi, ad esempio, non soltanto a documenti impaginati, bensì a documenti multimediali) dipendano dalle caratteristiche tecniche del formato scelto e delle modalità con cui il documento viene specificatamente formato, le considerazioni di cui alle sopracitate Linee guida saranno valutate allo scopo di incrementare l'interoperabilità e l'accessibilità dei documenti.
7. Proprio in considerazione dei precedenti 2, 4, 5 e 6, le PP.AA. possono includere nella valutazione di interoperabilità (cfr. §3.1) le considerazioni relative alla scelta dei formati dei file, delle modalità con cui i documenti vengono formati usando determinati formati (profili, impostazioni o configurazione del sistema di formazione del file, ecc.).
8. Nel caso di formati di file che permettano di scegliere i caratteri tipografici da utilizzare salvando tali scelte come parte del documento informatico stesso, qualora

i caratteri tipografici non facciano anch'essi parte del documento informatico, le PP.AA. adotteranno tutte le misure necessarie per la scelta di caratteri tipografici di uso comune, affinché la visualizzazione del documento sia il più possibile indipendente dai caratteri tipografici.
9. La raccomandazione di cui al punto 8 viene richiamata, nelle successive schede tecniche dei formati di file, mediante l'indicazione del capitolo sui caratteri tipografici, cioè con la semplice dicitura "cfr. §2.8", alla voce "Livello metadati".

2.1 Documenti impaginati

1. I formati "orientati alla pagina", o "impaginati", suddividono un documento informatico in unità separate chiamate *fogli*, in quanto il loro utilizzo primario è la carta stampata, come mezzo di origine (tramite scansione di documento cartaceo) ovvero di destinazione (tramite stampa). Non tutti i documenti impaginati sono però concepiti per la stampa, esibendo anzi caratteristiche tipiche del mondo digitale, quale la possibilità di allegare contenuti multimediali o anche di altro genere, così come di collegamenti di tipo URL a altre posizioni nella rete internet, intranet o in altro sistema di gestione dei documenti, in modo che l'utente possa accedervi tipicamente dopo aver attivato il collegamento di riferimento.
2. La pagina va inoltre considerata come un'entità virtuale — dunque non legata alla bidimensionalità e staticità di una superficie fisica di stampa. Con l'eccezione di alcuni formati nati ed evoluti esclusivamente nell'ambito della stampa tradizionale, e.g. il formato PostScript®, che dunque sono tecnicamente vincolati a rappresentare esclusivamente elementi direttamente stampabili (quali testi, disegni, immagini statiche), i formati più evoluti quali il PDF possono rappresentare dati multimediali, elementi interattivi o persino includere altri file all'interno di un documento impaginato (si legga la sezione corrispondente).
3. In formati di questo tipo viene descritta anche la "presentazione" (cioè la rappresentazione visuale) dei contenuti informativi all'interno delle pagine, utilizzando stili di visualizzazione di ipertesti (cfr. §2.2), immagini (raster o vettoriali, cfr. §2.6 e §2.7) e altre tecniche tipografiche, quali i caratteri tipografici (§0), che possono essere riferiti sia internamente nel medesimo file e che contiene dunque, al suo interno, un insieme completo o parziale di glifi.

PDF		FORMATO DI FILE
Nome completo	Portable Document Format	
Estensione/i	.pdf	
Magic number	%PDF	
Tipo MIME	application/pdf	
Sviluppato da	Adobe Systems	

Tipologia di standard	aperto (2.0)/proprietario (libero 1.7), estendibile, *de jure*
Livello metadati	4
Derivato da	Adobe® PostScript®
Revisione	2.0 (2017)
Riferimenti	Famiglie di standard **32000** e **19005** della ISO/IEC: • 32000-2:2017, *PDF v2.0* • 32000-1:2008, *PDF v1.7* • 19005-1:2005, *PDF/A-1* (v1.4) • 19005-2:2011, *PDF/A-2* (v1.7) • 19005-3:2012, *PDF/A-3* (v1.7) • Adobe, *Supplement to PDF v1.7 Extension 3*, ©2008 • Adobe, *Document management – PDF 1.7*, ©2008 • ISO 24517-1:2008, *PDF/E-1* (v1.6) • ISO 15930-1:2001, *PDF/X-1 e PDF/X-1a* (v1.4) • ISO 15930-8:2010, *PDF/X-5* (v1.6) • ISO 14289-1:2014, *PDF/UA-1* (v1.4) • ISO/CD 14289-2, *PDF/UA-2* (v2.0) • ISO 16612-2:2010, *PDF/VT-1 e PDF/VT-2* (PDF/X-4 e /X-5) • ISO/CD 16612-3, *PDF/VT-3* (PDF/X-6) (v2.0)
Conservazione	Sì, solo profili PDF/A e PDF/B; altrimenti, cfr. §2.8
Racc. per la lettura	Generico con riconoscimento obbligatorio (v1.*x*).
Racc. per la scrittura	Raccomandata: v1.7+. Obbligatoria: v1.4+. Profili raccomandati (leggere raccomandazioni più sotto): PDF/A-2a, PDF/A-2u, PDF/A-2b, PDF/A-1a, PDF/A-1b. Profili PDF/A e PDF/B adatti alla conservazione.

4. Il formato PDF è il "principe" dei formati per documenti impaginati. Nato come formato proprietario (sviluppato da Adobe Corporation), è stato rapidamente adottato come standard *de facto* per la produzione digitale del cartaceo; nel 2008 è diventato uno standard *de iure*, con il nome di ISO 32000-1 (PDF versione 1.7); successivamente rivisto nel 2017, con il nome di ISO 32000-2 (PDF versione 2.0). Il formato PDF è estremamente versatile in quanto è stato continuamente aggiornato nel tempo aggiungendo nuove funzionalità. Allo stato attuale (PDF versione 2.0) è possibile includere nel documento evidenze di vario tipo (incluse immagini, audio, video, modelli tridimensionali animati), imbustare all'interno del file di qualsivoglia formato sotto forma di "allegati", sovraimprimere graficamente alle pagine moduli interattivi che permettono all'utente compilazione libera o vincolata (per poi salvarne un'istanza con i campi riempiti). È anche possibile proteggere un documento PDF con password o firme elettroniche e usare la crittografia per

limitarne la lettura o la modifica, anche se tale protezione è aggirabile con appositi strumenti software.
5. L'apposizione di firme e sigilli elettronici su documenti in formato PDF è effettuata mediante la busta crittografica PAdES (cfr. §2.16), mentre il servizio fiduciario elettronico costituto dalla convalida di documenti in formato PAdES è sancito dalla normativa comunitaria.
6. Il PDF è uno standard "modulare", nel senso che sono stati definiti profili diversi che introducono insiemi aggiuntivi di nuove funzionalità (quali quelle sopra elencate), ovvero vincolano alcune di esse allo scopo di migliorare l'interoperabilità del documento PDF in specifici ambiti. I profili possono prevedere sotto-profili con ulteriori specifiche (additive o vincolanti), chiamati livelli di conformità. Si noti inoltre che un file PDF può essere conforme a più livelli (profili o sotto-profili) contemporaneamente. I profili e sotto-profili ufficialmente riconosciuti sono i seguenti:

- **PDF/A** (*archival*) — Profilo particolarmente adatto alla creazione di documenti di cui deve essere garantita la leggibilità in caso di archiviazione a lungo termine e conservazione. Si divide in due sotto-profili:
  - **PDF/A-1** — Basato su PDF versione 1.4, impedisce al file di contenere (graficamente o come allegati) niente altro che testi, ipertesti, immagini *raster* o vettoriali; sono in particolar modo vietati i moduli con contenuti variabili (e.g. codice Java eseguibili), Infine, il formato contenere al suo interno tutti i caratteri tipografici utilizzati.
    - **PDF/A-1a** (*accessible*) — Specifica "forte" del sotto-profilo PDF/A-1, ove ad ogni contenuto non testuale del file è garantita piena coerenza semantica e accessibilità (p.es. ogni immagine deve avere un commento, ed ogni glifo un codice UNICODE); questo consente non solo una visualizzazione del file a lungo termine, ma anche un suo utilizzo da parte di *parser* e di lettori del documento per persone diversamente abili.
    - **PDF/A-1b** (*basic*) — Specifica "debole" del sotto-profilo PDF/A-1, ove non è richiesta la presenza di dati semantici come in PDF/A-1a.
  - **PDF/A-2** — Basato su PDF versione 1.7, è dotato di tre livelli di conformità:
    - **PDF/A-2a** — analogo a PDF/A-1a.
    - **PDF/A-2b** — analogo a PDF/A-1b.
    - **PDF/A-2u** — analogo a PDF/A-2b, ma con il vincolo aggiuntivo che tutti i caratteri impiegati devono essere codificati in base alla mappatura UNICODE.

    - **PDF/A-3** — Sotto-profilo poco usato che aggiunge, rispetto a PDF/A-2, la possibilità di allegare file di qualsiasi tipo.
- **PDF/E** (*engineering*) — Profilo dedicato all'inclusione di disegni e altri dati tecnici nel PDF (versione 1.6), quali informazioni geografiche, e modelli grafici tridimensionali interattivi.
- **PDF/H** (*health*) — Inclusione di dati sanitari (sotto forma di XML) quali referti, sondaggi, testi di laboratorio o altri sistemi diagnostici.
- **PDF/X** (*graphics exchange*) — Interscambio professionale di contenuti grafici (sia *raster* che vettoriali, cfr. §2.6 e §2.7 rispettivamente)
    - **PDF/X-1** — Sotto-profilo basato su PDF versione 1.3.
        - **PDF/X-1a** — Variante in cui devono essere incluse tutte le fonti tipografiche, mentre i colori sono codificati in spazio-colore CMYK o come colori *spot*[10].
    - **PDF/X-3** — Sotto-profilo (basato su PDF versione 1.3) che include capacità colorimetriche avanzate, come ad esempio ammettere, oltre a colori *spot*, spazi-colore CMYK, CIELAB, RGB (calibrati o meno), profili ICC.
    - **PDF/X-4** — Ulteriore evoluzione del sotto-profilo PDF/X-3 (basato su PDF versione 1.4) che include anche le trasparenze.
    - **PDF/X-5** — Ulteriore evoluzione del sotto-profilo PDF/X-5 (basato su PDF versione 1.6), con una ulteriore distinzione: **PDF/X-5g**, **PDF/X-5pg** e **PDF/X-5n**.
- **PDF/UA** (*universal accessibility*) — Profilo dedicato all'accessibilità universale, che comprende ad esempio l'ordine con cui devono comparire commenti e note, le didascalie per i contenuti extra meno accessibili (ad es. foto o video), i requisiti per permettere a tecnologie di accessibilità di gestire tutte le parti del documento (incluse quelle criptate).
- **PDF/VT** (*variable and transactional printing*) — Profilo dedicato a PDF che andranno stampati con parti variabili da tiratura a tiratura (ad es. il numero seriale progressivo in ogni esemplare a tiratura limitata).
    - **PDF/VT-1** — Variante in cui tutte le parti sono auto-contenute in un singolo file PDF (compatibile con PDF/X-4).
    - **PDF/VT-2** — Variante in cui è contemplato l'uso di elementi grafici contenuti in file esterni (compatibile con PDF/X-5)

[10] Un colore *spot* è indicato digitalmente da un semplice codice corrispondente, nei plotter e stampanti professionali, ad un particolare inchiostro che viene inserito appositamente. I colori spot possono prevedere qualità colorimetriche non visibili digitalmente (p.es. colori lucidi, opachi, metallizzati, perlati, ruvidi), così come essere associati, in particolari processi di stampa, a lavorazioni completamente diverse nella zona dell'impaginato digitalmente indicata come da stampare con un dato color *spot*.

- **PDF/VT-2s** — Ulteriore variante in cui sono ammessi contenuti presi da flussi digitali esterni al PDF stesso (p.es. dati provenienti da un sensore).

7. Come riportato nella scheda tecnica del formato PDF, a prescindere dalla tipologia di contenuto (e quindi dall'eventuale conformità con altri profili), la produzione di file in formato PDF privilegia, ove disponibile, la più recente versione conforme allo ISO 32000-2, adottando preferenzialmente i suoi i profili di accessibilità, con il seguente ordine di preferenza: PDF/A-2a, PDF/A-2u, PDF/A-2b. Qualora non siano disponibili strumenti per la produzione di documenti PDF conformi alla versione 2.0, si utilizza la versione più recente disponibile, preferendo le versioni conformi allo ISO 32000-1 (PDF versione 1.7 e successive), in nessun caso inferiori al PDF versione 1.4, adottando i profili di accessibilità con il seguente ordine di preferenza: PDF/A-1a, PDF/A-1b. La ragione principale di tale ordine di preferenza risiede nella maggiore versatilità del PDF versione 1.7 e successive nel:

- includere documenti (anche non in formato PDF), come *allegati* al PDF stesso;
- permettere l'utilizzo di campi modulo senza incidere sulla validità di firme o sigilli elettronici senza invalidare l'integrità del documento stesso, né tantomeno delle firme o sigilli elettronici PAdES eventualmente apposti;

8. L'obbligo di cui al precedente capoverso può essere disatteso per i documenti PDF prodotti per essere specificatamente modificati ovvero "compilabili" in un momento successivo alla loro produzione, come ad "modelli" o *template* per procedimenti amministrativi di varia natura. Tuttavia, le PP.AA. producono il documento compilabile e modificabile in formato PDF, usando versioni e profili adeguati, tenendo conto che una volta compilati o redatti in forma definitiva, tali file costituiranno documenti informatici (nel senso della rilevanza degli atti, fatti o dati giuridicamente rilevanti contenuti nella versione modificata o compilata). Si raccomanda, ad esempio, di utilizzare per file PDF compilabili o modificabili, solamente caratteri tipografici tra quelli interoperabili "standard" definiti nel §2.8.
9. I documenti che sono immodificabili salvo per la compilazione di campi vuoti o l'apposizione di firme o sigilli elettronici devono essere prodotti in formato PDF versione almeno pari a 1.7, sfruttando i campi modulo e la loro robustezza, come già esposto al punto 7.

WORD® 2007		FORMATO DI FILE
Nome completo	WordProcessingML OOXML Extension	
Estensione/i	.docx, .dotx	
Specializzazione di	XML imbustato dentro ZIP	
Tipo MIME	application/vnd.openxmlformats-officedocument.wordprocessingml.document application/vnd.openxmlformats-officedocument.wordprocessingml.template	

Sviluppato da	Microsoft Corporation; ISO; ECMA
Tipologia di standard	proprietario (libero), estendibile, *de facto*, *testuale*
Livello metadati	3
Derivato da	Office Open XML; Microsoft® *Word* ®
Revisione	11.1 (2018)
Riferimenti	• Microsoft, *Word extensions to OOXML (.docx) file format* v11.1 (2018) • officeopenxml.com, *Anatomy of a WordProcessingML file*
Conservazione	Sì, solo profilo Strict; cfr. §2.8
Racc. per la lettura	Generico con riconoscimento obbligatorio
Racc. per la scrittura	Vedasi capoversi 10 e 11 per la conservazione.

9. La suite di applicativi Microsoft® *Office*®, dalla versione 2007 in poi, utilizza un formato di file unico per i suoi applicativi principali, chiamato Open Office XML (OOML, descritto più genericamente in §2.5). Il formato consiste in un pacchetto di file, suddiviso in più cartelle, imbustato e compresso con un algoritmo ZIP e presentato come un unico file. I file all'interno del pacchetto compresso sono prevalentemente in formato XML e utilizzano il dialetto WordprocessingML, riservato ai documenti di videoscrittura elaborati dall'applicativo *Word*® per definire l'intera struttura e il contenuto del documento. Eventuali documenti allegati (p.es. immagini, video, audio o altri file) sono inseriti, all'interno di opportune cartelle, nel loro formato nativo. L'estensione del documento compresso per gli impaginati normali (*senza macro* attive) è .docx, mentre altri tipi di documenti (p.es. modelli di documento, impaginati con *macro* attive) sono indicati semplicemente con estensioni diverse.
10. Come per altri formati basati su OOXML, si consiglia la produzione di documenti con il profilo Strict, che è più restrittivo ma consente di eliminare alcune estensioni "proprietarie" che possono ridurre l'interoperabilità del formato stesso.
11. Il documento è adatto alla conservazione solo se:

- sono utilizzati esclusivamente caratteri tipografici "standard" (cfr. §2.8 capoverso 2),
- è privo di contenuti dinamici ad eccezione di campi compilabili o campi-firma,
- è privo di contenuti audiovisivi (suoni, video),
- eventuali immagini o altri contenuti multimediali sono contenute direttamente nel documento e non mediante collegamenti a file esterni al documento.

Si consiglia inoltre di effettuare un controllo sull'intera accessibilità del documento.

MS-DOC		FORMATO DI FILE
Nome completo	Microsoft® *Word* ® Binary File Format	
Estensione/i	.doc, .dot	
Magic number	0xD0CF11E0A1B11AE1	
Tipo MIME	application/msword	
Sviluppato da	Microsoft Corporation	
Tipologia di standard	proprietario (libero), estendibile, *de facto*, binario, deprecato	
Livello metadati	3	
Derivato da	Microsoft® Compound File Binary format; Corel® *WordPerfect*™	
Revisione	8.1 (2018)	
Riferimenti	• Microsoft, *[MS-DOC]: Word (.doc) binary file format* v8.1 (2018)	
Conservazione	No; cfr. §2.8	
Racc. per la lettura	Obbligatorio con riversamento raccomandato	
Racc. per la scrittura	Sconsigliato	

12. Il formato binario *Word*,® utilizzato come formato principale fino alle versione 2003 da codesto applicativo di videoscrittura, è una variante specializzata del formato Compound File Binary, proprietario di Microsoft Corporation (cfr. §2.5); utilizza l'estensione .doc per il documenti impaginati normale, e la cambia in caso di varianti, quali ad esempio i modelli di documento (estensione .dot). Il formato contiene in un unico file i metadati, i contenuti testuali, ipertestuali e gli allegati dell'impaginato, senza offrire efficaci meccanismi di controllo dell'integrità. Inoltre, le revisioni al documento vengono in generale salvate come modifiche differenziali in coda al file, contribuendo ad ingrandire la dimensione del file e, al tempo stesso, renderne più complessa l'apertura e l'interpretazione da parte degli applicativi di videoscrittura. Per questa mancanza di robustezza (soprattutto in caso di documenti di dimensioni molto grandi a causa di contenuti multimediali allegati), l'azienda proprietaria del formato decide di cambiare strategia, adottando un nuovo formato per tutti i documenti della suite applicativa *Office*,® a partire dalla versione 2007, anche se il "formato .doc" è pienamente supportato dalle nuove versioni. In caso di produzione di nuovi documenti impaginati tramite *Word*® si raccomanda l'uso del nuovo, sopra descritto formato basato su Open Office XML.

ODT		FORMATO CONTENITORE
Nome completo	Open Document Text	
Estensione/i	.odt	
Specializzazione di	XML imbustato dentro ZIP	
Tipo MIME	application/vnd.oasis.opendocument.text	
Sviluppato da	Organization for the Advancement of Structured Information Standards	

Tipologia di standard	aperto, estendibile, *de iure*, binario
Livello metadati	3
Derivato da	—
Revisione	1.2 (2015)
Riferimenti	Famiglia di standard **26300** della ISO/IEC: • ISO/IEC 26300-1:2015, *ODF for Office Applications v1.2 – Part 1: OpenDocument Schema* • ISO/IEC 26300-3:2015, *ODF for Office Applications v1.2 – Part 3: Packages* • OASIS, *Open Document Format for Office Applications (OpenDocument)*, v1.2 (2015)
Conservazione	Sì; cfr. §2.8
Racc. per la lettura	Generico con riconoscimento obbligatorio
Racc. per la scrittura	Fortemente raccomandato

13. Il formato OpenDocument Text (ODT) è una particolare implementazione del più generale formato OpenDocument (cfr. §2.5). Esso è il formato di default usato dall'applicazione di videoscrittura *Writer* della suite open source *LibreOffice*, ma è ampiamente supportato da molti altri applicativi di videoscrittura. Come per l'Open Office XML, anche tale formato è costituito da una busta compressa ZIP contenente un pacchetto di file che descrive la composizione del documento, cui componenti principali sono in formato XML.

RICHTEXT		FORMATO DI FILE
Nome completo	Rich Text Format	
Estensione/i	.rtf	
Magic number	{\rtf1	
Tipo MIME	text/rtf, application/rtf	
Sviluppato da	Microsoft	
Tipologia di standard	proprietario (libero), *de facto*, binario, deprecato	
Livello metadati	1	
Derivato da	Rich-Text; Enriched Text	
Revisione	1.9.1 (2008)	
Riferimenti	• Microsoft, *Word 2007: Rich Text Format (RTF) Specification* v1.9.1 (2008) • RFC-1521	
Conservazione	No; cfr. §2.8	
Racc. per la lettura	Generico con riconoscimento consigliato	
Racc. per la scrittura	Non raccomandato	

14. Rich Text Format (RTF) è un formato di proprietà di Microsoft utilizzato come formati interoperabile di documenti impaginati. Il formato supporta un numero molto ridotto di caratteristiche grafiche e tipografiche, e una possibilità limitata relativamente a ipertesti e allegati multimediali, a fronte di una semplicità strutturale conforme con lo scopo della sua introduzione. Nonostante sia ancora supportato dai principali applicativi di videoscrittura (inclusi quelli installati di default nei principali sistemi operativi), ma seguito della standardizzazione e apertura di formati ben più evoluti e completi (cfr. quelli basati su OOXML e OpenDocument) e, se ne sconsiglia l'uso per la produzione di nuovi documenti.

ePub	Formato di file
Nome completo	EPUB
Estensione/i	.epub
Specializzazione di	HTML, CSS, XML, SVG, … imbustati dentro ZIP
Tipo MIME	application/epub+zip
Sviluppato da	International Digital Publishing Forum
Tipologia di standard	aperto, estendibile, *de facto*
Livello metadati	3; cfr. §2.8
Derivato da	Open eBook publication structure; XHTML; CSS
Revisione	3.1 (2017)
Riferimenti	Famiglia di standard **30135** della ISO/IEC: • ISO/IEC 30135-1:2014, *EPUB Part 1: EPUB3 overview* • ISO/IEC 30135-3:2014, *EPUB Part 3:content documents* • ISO/IEC 30135-5:2014, *EPUB Part 5: media overlay* • www.idpf.org/epub
Conservazione	No; cfr. §2.8
Racc. per la lettura	Generico con riconoscimento consigliato
Racc. per la scrittura	Raccomandato nell'editoria digitale

15. Il formato EPUB è costituito da un pacchetto di file che descrivono, complessivamente, una documento paginato comprensivo dei suoi allegati multimediali, compressi con compressione ZIP in un unico file con estensione .epub. I documenti prediletti da questo formato si chiamano anche eBook, in quanto si prestano ad una visualizzazione prevalentemente elettroica su dispositivi con diverse

capacità di visualizzazione grafica, anche notevolmente ridotta, come ad esempio *smartphone* e *tablet* (inclusi gli '*eBook reader*' specializzati).[11]
All'interno di questo pacchetto, i metadati interni del documento sono contenuti in file XML, il contenuto testuale in uno o più file HTML, le sue impostazioni tipografiche in uno o più file CSS, e gli eventuali elementi grafici in file immagini di vari formati aperti (PNG, SVG, ecc.). La versatilità grafica dell'*eBook* è ottentuta da questo formato sfruttando le caratteristiche di adattamento ed "elasticità" dei contenuti HTML dotati di opportuni "stili" CSS (cfr. §2.2) basati su profili di dispositivo apposito.

INDESIGNML	FORMATO DI FILE
Nome completo	Adobe® *InDesign*® Markup Language
Estensione/i	.idml
Specializzazione di	XML imbustato dentro ZIP
Tipo MIME	application/x-indesign+xml
Sviluppato da	Adobe Systems
Tipologia di standard	proprietario (libero), *de facto*, testuale
Livello metadati	4
Derivato da	Adobe® *InDesign*® file format (.indd)
Revisione	8.0 (2012)
Riferimenti	• Adobe, *IDML File Format Specification*, v8.0 (2012) • Adobe, *InDesign*® Developer Documentation
Conservazione	No; cfr. §2.8
Racc. per la lettura	Specifico; non raccomandato se non in editoria digitale
Racc. per la scrittura	Specifico; raccomandato (al posto di .indd) nel settore dell'editoria digitale *InDesign*

16. Adobe® *InDesign*® è uno degli applicativi più diffusi per l'editing grafico, ma è un applicativo commerciale non aperto, come il formato nativo e binario per i suoi impaginati (estensione .indd). Tale formato è inoltre altamente instabile in quanto Adobe, che lo mantiene, lo cambia in continuazione da una variante all'altra del software. Per ridurre i rischi di interoperabilità dovuti all'estrema variabilità del formato nativo ".indd", Adobe ha introdotto una variante del formato basato su XML e chiamato *InDesign*® Markup Language (estensione .idml). Tale versione, oltre ad essere testuale ed estendibile, è anche parzialmente pubblicata sul sito Adobe, ed

[11] Ad esempio lo stesso *eBook* visualizzato su due dispositivi elettronici con display tecnicamente diversi (un '*e-paper*' monocromatico da 800×600 punti a 167dpi, rispetto a un LCD a colori da 1920×1080 — cfr. §2.6) potrebbe essere visualizzato con caratteri tipografici simili (magari di dimensioni diverse) e con una diversa distribuzione di caratteri su ogni linea allo scopo di rendere l'esperienza di lettura più agevole e la fruibilità più simile a quella di un libro su carta stampata.

è perciò consigliato come il formato d'elezione –al posto del formato nativo– per tutte le organizzazioni che archivino impaginati creati con *InDesign*®.

POSTSCRIPT™	FORMATO DI FILE
Nome completo	PostScript®
Estensione/i	.ps
Magic number	%!PS
Tipo MIME	application/postscript
Sviluppato da	Adobe Systems
Tipologia di standard	proprietario (libero), *de facto*, binario, deprecato
Livello metadati	1
Derivato da	Lisp
Revisione	3 (1997)
Riferimenti	• Adobe, *PostScript® Language reference*, 3rd ed. (1999) • Adobe, *PostScript® Language reference supplement: Adobe® PostScript® 3™ version 3010 and 3011 Product Supplement*, 30 August 1999
Conservazione	No
Racc. per la lettura	Raccomandato per l'editoria digitale
Racc. per la scrittura	Sconsigliato

17. Il formato PostScript™ deve il suo nome all'omonimo linguaggio di descrizione di pagine professionale. Come per i formati contenenti codici scritti in un linguaggio di programmazione, questo formato di file contiene un elenco ordinato di istruzioni in tale linguaggio PostScript™ (linguaggio di tipo pseudo-binario) che, eseguite una dopo l'altra da una stampante o altro dispositivo capace di interpretarle, hanno come effetto la stampa di un documento esattamente identico. Il PostScript™ è stato per anni uno dei linguaggi di riferimento per il trasporto e la conservazione di documenti finalizzati alla stampa digitale, sia di tipo testuale che grafico. La versatilità del formato è anche parte del motivo per cui lo standard PDF è basato su PostScript™. Ai formati per la produzione di documenti impaginati o altri tipi di stampa digitale, è tuttavia richiesto di supportare anche ipertesti e contenuti non strettamente stampabili (suono, video ,ecc.), per i quali tale formato, risultando dunque inadeguato, diviene obsoleto. Nonostante la grande varietà di documenti attualmente presenti salvati in questo formato (per il quale la capacità di lettura è ancora fortemente raccomanda alle organizzazioni di settore tecnico-editoriale), si sconsiglia di produrne degli altri, preferendovi formati più moderni. Gli applicativi dedicati alla formazione del documenti in tali formati al posto del PostScript™, così come gli applicativi di riversamento dal PostScript™ in questi formati devono tuttavia essere adeguatamente configurati affinché i documenti mantengano le caratteristiche di riproducibilità e qualità che il formato PostScript™ garantisce per

sua natura, permettendo al contempo di includere funzionalità moderne nel documento, quali contenuti ipertestuali e multimediali. Esempi di tali funzionalità sono la rappresentazione delle immagini in modalità vettoriale o, qualora le immagini siano di tipo raster, garantire qualità equivalente (ad esempio mediante algoritmi di compressione senza perdita); cfr. §2.6 e §2.7.

LaTeX		Formato di file
Nome completo	LaTeX	
Estensione/i	.tex	
Specializzazione di	TeX	
Tipo MIME	application/x-tex	
Sviluppato da	comunità open source	
Tipologia di standard	aperto (licenza LPPL), estendibile, *de facto*, testuale	
Livello metadati	1	
Derivato da	TeX	
Revisione	2ε	
Riferimenti	• www.latex-project.org • github.com/latex3 • ctan.org (Comprehensive TeX Archive Network)	
Conservazione	No; cfr. §2.8	
Racc. per la lettura	Specifico; raccomandato per testi tecnico-scientifici	
Racc. per la scrittura	Specifico; nessuna raccomandazione	

18. LaTeX è la specializzazione più comunemente più usata del linguaggio di impaginazione testi denominato TeX, inventato nel 1979 da Donald Knuth. La versione attualmente in uso è la 2ε (denominata LaTeX2ε). In realtà, l'albero "glottologico" del TeX comprende dialetti derivati direttamente dal linguaggio-madre TeX, così come dal LaTeX stesso, come ad esempio *AMS*-LaTeX.

Come suggerisce il nome stesso, LaTeX è particolarmente efficace per la produzione di pubblicazioni tecnico-scientifiche pronte per la stampa e conformi a norme redazionali e tipografiche professionali. Ciò che ne ha determinato particolare diffusione nella comunità scientifica internazionale (nonostante la sua obsolescenza rispetto sia a linguaggi per ipertesti più evoluti come l'XML e a formati di documenti impaginati "WYSIWYG",[12] che permettono una visualizzazione diretta dei contenuti) sta, in prima istanza, nella capacità –innata nel TeX e contemporaneamente potenziata e semplificata nel LaTeX– di descrivere complesse formule matematiche, la cui tipografia è rigorosamente rappresentata con leggibilità superiore ad applicativi di videoscrittura da "ufficio". Il linguaggio, di per sé, si limita

[12] Acronimo dall'inglese «*what you see is what you get*» (parafrasabile in «esattamente così come lo vedi»).

a descrivere la distribuzione di testi e ipertesti (immagini, tabelle, formule/equazioni, bibliografia, note, così come i loro riferimenti) all'interno del documento, contornati da *tag* che si riferiscono a *template* esterni contenenti le regole stilistiche e tipografiche da rispettare (come ad esempio le fonti, cfr. §2.8). Il documento LaTeX è, in questo senso, analogo al codice sorgente di un linguaggio di programmazione o scripting (cfr. §2.15): ha bisogno di essere *compilato* per produrre un documento informatico consultabile e completamente impaginato, che si raccomanda sia prodotto sempre nel formato PDF o comunque –ove ciò non sia possibile– in uno dei formati raccomandati in questo paragrafo per la produzione di documenti impaginati.
È attualmente in sviluppo una revisione maggiore delle specifiche, denominata LaTeX3, che modifica diversi aspetti (sia dal punto di vista lessicale che grafico), allo scopo di attualizzare questo linguaggio e integrarlo con le tecnologie allo stato dell'arte nel campo della tipografia, dell'editoria e della multimedialità che, semplicemente, non esistevano negli anni '80 e la cui integrazione nel TeX e nei dialetti derivati è solo parzialmente avvenuta, non senza numerose difficoltà tecniche.

2.1.1 Raccomandazioni per la produzione di documenti

1. 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).
2. Le caratteristiche avanzate quali l'apposizione di firme e sigilli elettronici anche multipli (si veda il formato PAdES, §2.16), 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.
3. 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

puramente "virtuali" quali quelli delle suite collaborative di fornitori di servizi in Cloud qualificati.
4. Per applicazioni specifiche, come l'editoria, la grafica o le pubblicazioni tecnico-scientifiche, possono essere raccomandabili anche altri formati definiti in questa sezione purché, in fase di conservazione, si valuti sempre l'interoperabilità del documento finito (a scopo di distribuzione attraverso vari mezzi di comunicazione) e della sua sorgente così (come generata da applicativi di videoscrittura o altro), in quanto tali requisiti potrebbero portare a differenti raccomandazioni di formato.

2.2 Ipertesti

1. Rientrano in questa categoria tutti i documenti, prevalentemente testuali, che contengono ipertesti, quali ad esempio riferimenti ad oggetti esterni (di tipo URI, URL o altro), ovvero codici di *markup* per rappresentare digitalmente insieme astratti di dati e loro ontologie.

XML	FORMATO DI FILE
Nome completo	Extensible Markup Language
Estensione/i	.xml
Magic number	<?xml 0x20
Tipo MIME	application/xml, text/xml
Sviluppato da	World Wide Web Consortium
Tipologia di standard	aperto, estendibile, retrocompatibile, *de iure*, testuale
Livello metadati	4
Derivato da	SGML
Revisione	1.0, 5ª edizione
Riferimenti	• W3C Recommendation *XML 1.0 (5th Ed.)*, 2013 • www.w3.org/standards/xml • validator.w3.org
Conservazione	Sì, se conservato insieme a un *XML Schema*
Racc. per la lettura	Generico; subordinato ad eventuali, ulteriori obblighi o raccomandazioni di conformità con schemi/dialetti.
Racc. per la scrittura	Generico; subordinato ad eventuali, ulteriori obblighi o raccomandazioni di conformità con schemi/dialetti.

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

- possono essere definite in continuazione nuove etichette;

- più dizionari di etichette possono essere usati con un approccio modulare all'interno del medesimo file XML mediante l'utilizzo dei *namespace* ("spazi di nomi") sui nomi delle etichette;
- è possibile definire, in XML stesso, sintassi specializzate mediante schemi (.xsd) o dialetti (.dtd);
- un medesimo documento XML (come file unico ovvero pacchetto costituito da più file XML) può essere logicamente composto da parti distinte, ciascuna delle quali utilizza etichette e regole sintattiche definite da diversi *namespace* o schemi, mentre le diverse parti possono referenziarsi l'un l'altra sfruttando diversi meccanismi sintattici (p.es. *XQuery* e *XPath*).

3. Numerosi formati di file oggetto del presente Allegato che utilizzano il linguaggio XML impiegano un dialetto specializzato al loro contenuto; essi sono:

- OpenDocument e Microsoft® OOXML (cfr. § 2.1 e §2.5),
- XHTML, XSD, XSL, XSLT, MathML (§2.2),
- SVG (§2.7),
- tutti i formati descritti in §2.14 (FatturaPA, CDA2, asserzione SPID, …),
- MusicXML (§2.9),
- IMSC1, TTML, e EBU-TT (§ 2.11)
- alcuni file obbligatori nei pacchetti IMF e DCP, nonché ACESclip (§2.12),
- KDM e firme elettroniche XAdES (§2.16).

4. ***Nota Bene***: Alcuni dei formati basati su XML sopraelencati mantengono l'estensione .xml del linguaggio madre, altri usano le proprie (e.g. .html/.htm, .svg, .ttml, .kdm). Si raccomanda che le Pubbliche Amministrazioni che producano documenti informatici in qualunque formato basato su XML non descritto dal presente Allegato ma che utilizzasse (in base alle proprie linee guida, specifiche tecniche, raccomandazioni o *best practices*) un'estensione "propria" diversa da .xml, laddove fosse consentito dall'uso applicativo delle applicazioni che elaborano file in tali formati, di "appendere" l'estensione .xml alla propria. Questo accorgimento manterrà ed esporrà l'estensione propria del formato a livello di UI, consentendo ai sistemi operativi privi di applicazioni specifiche per interpretare tali formati di aprire e visualizzare tali file come fossero normali file XML.
Esistono moltissime altre estensioni di XML che, pur non descritte qui, possono essere utilizzate e integrate in documenti XML per mezzo dell'estendibilità del *namespace*.

HTML		FORMATO DI FILE
Nome completo	Hypertext Markup Language	
Estensione/i	.html, .htm	
Magic number	<!DOCTYPE 0x20; <head>	
Tipo MIME	text/html	

Sviluppato da	World Wide Web Consortium
Tipologia di standard	aperto, estendibile, *de iure*, testuale
Livello metadati	2
Derivato da	XML
Revisione	5.2
Riferimenti	• W3C Recommendation *HTML 5.2*, 2017 • validator.w3.org • W3C Recommendation *XML 1.0 (5th Ed.)*, 2013
Conservazione	Sì, se conservato insieme al/i CSS; cfr. §2.8
Racc. per la lettura	Generico con riconoscimento obbligatorio
Racc. per la scrittura	Specifico; raccomandato HTML5 per contenuti web

5. Un particolare dialetto di XML è HTML (ufficialmente codificato come XML puro a partire dalle sue versioni "XHTML" e "HTML5"), che serve a rappresentare il contenuto di pagine web. In questo caso HTML (.htm, .html) è affiancato dal CSS, linguaggio specifico per descrivere i fogli di stile che trasformano i componenti logici di una pagina in elementi grafici (la cosiddetta "presentazione" della pagina).
6. Il CSS può essere iniettato direttamente all'interno dell'HTML, oppure venire referenziato da esso in dei file esterni (estensione .css).
7. Sia XML, che HTML, che CSS sono tutti standard del W3C. Più in generale, una pagina web è un esempio di pacchetto di file (cfr. §1.1.2) costituito, come minimo, da un solo indice HTML, più eventualmente altri file in vari formati (anche non inclusi in questo Allegato), quali ad esempio HTML, CSS, WOFF, JavaScript, ecc.

XHTML		FORMATO DI FILE
Nome completo	Extensible Hypertext Markup Language	
Estensione/i	.xhtml, .html	
Specializzazione di	XML	
Tipo MIME	application/xhtml+xml	
Sviluppato da	World Wide Web Consortium	
Tipologia di standard	aperto, estendibile, *de iure*, testuale	
Livello metadati	2	
Derivato da	HTML	
Revisione	1.1	
Riferimenti	• W3C Recommendation *XHTML™ Basic 1.1 (2nd Ed.)*, 2010 • W3C Recommendation *XHTML™ 1.0 (2nd Ed.)*, 2018 • validator.w3.org • W3C Recommendation *XML 1.0 (5th Ed.)*, 2013	

Conservazione	Solo se conservato insieme al/i CSS
Racc. per la lettura	Generico con obbligo di riconoscimento
Racc. per la scrittura	Specifico; raccomandato HTML5 per contenuti web

8. L'HTML, originariamente (e fino a HTML 4.01), non era una specializzazione di XML; prima di introdurre la versione 5, è stato fatto dal W3C un tentativo di standardizzazione intermedio che ha portato ad un'altra versione di HTML, chiamato XHTML (a sua volta ramificato in due dialetti distinti –*Strict* e *Transitional*– sono i più diffusi, e in diverse versioni). Data la sua diffusione tale linguaggio (che condivide le estensioni di file con le versioni ufficiali di HTML) è stato incluso in questo elenco.

XSD		FORMATO DI FILE
Nome completo	XML Schema Definition	
Estensione/i	.xsd	
Specializzazione di	XML	
Tipo MIME	application/xml	
Sviluppato da	World Wide Web Consortium	
Tipologia di standard	aperto, estendibile, *de iure*, testuale	
Livello metadati	4	
Derivato da	XML	
Revisione	2.0	
Riferimenti	• W3C Recommendation *XSL Schema Part 0: Primer, 2nd Ed.*, 2004 • W3C Recommendation *XSL Schema Part 1: Structures, 2nd Ed.*, 2004 • W3C Recommendation *XSL Schema Part 1: Datatypes, 2nd Ed.*, 2004	
Conservazione	Sì	
Racc. per la lettura	Speciale; per la convalida di documenti in XML	
Racc. per la scrittura	Speciale; per la condivisione di sintassi e dialetti XML	

9. Questo formato è in grado di descrivere, sempre in linguaggio XML la sintassi e la grammatica associata ad un particolare schema, che esso stesso definisce. Un documento XML può perciò essere convalidato in maniera automatica rispetto ad un dato schema, per verificare se ne rispetta tutti i criteri sintattici. Analogamente, un file XSD può essere usato per produrre in XML una struttura di dati o, più in generale, una procedura. In entrambe i casi (convalida o produzione di un documento XML rispetto ad uno schema), questa architettura consente di riutilizzare lo stesso algoritmo o applicativo, ma con la flessibilità aggiuntiva di poter

cambiare le regole sintattiche e grammaticali in automatico non appena viene cambiato lo schema di riferimento su file XSD.

XSL	FORMATO DI FILE
Nome completo	Extensible Stylesheet Language
Estensione/i	.xsl
Specializzazione di	XML (*namespace* xsl)
Tipo MIME	text/xsl
Sviluppato da	World Wide Web Consortium
Tipologia di standard	aperto, estendibile, *de iure*, testuale
Livello metadati	4
Derivato da	XML
Revisione	2.0
Riferimenti	• W3C Working Draft *XSL Requirements*, v2.0, 2008 • W3C Recommendation *XSL Requirements*, v1.1, 2006 • www.w3.org/Style/XSL/
Conservazione	Sì
Racc. per la lettura	Generale; raccomandato per la visualizzazione di XML
Racc. per la scrittura	Speciale; raccomandato per la costruzione di presentazioni grafiche di documenti in formato XML

10. Un documento esistente in formato XML può essere trasformato in un altro documento, sia esso in XML o in qualunque altro formato, specificando le regole di traduzione mediante trasformazioni descritte –sempre in linguaggio XML– e raccolte in:

- File per la descrizione del linguaggio di stile (in caso si tratti di un linguaggio diverso da XML) o del dialetto (in caso si tratti di una specializzazione di XML) — in XSL (.xsl).

XSLT	FORMATO DI FILE
Nome completo	Extensible Stylesheet Language Transformations
Estensione/i	.xslt
Specializzazione di	XML (*namespace* xsl)
Tipo MIME	application/xslt+xml, text/xml
Sviluppato da	World Wide Web Consortium
Tipologia di standard	aperto, estendibile, *de iure*, testuale
Livello metadati	3
Derivato da	DSSSL
Revisione	2.0

Riferimenti	• W3C Recommendation *XSL Transformations (XSLT) v2.0*, 2007
Conservazione	Sì
Racc. per la lettura	Generale; raccomandato per la visualizzazione di XML
Racc. per la scrittura	Speciale; raccomandato per la costruzione di presentazioni grafiche di documenti in formato XML

- File per la descrizione delle trasformazioni del linguaggio di stile, che permette regole più complesse che, ad esempio, possono includere il traversamento della struttura XML del documento di partenza (tramite i sopracitati metodi *xQuery* e *xPath*) —XSLT (.xslt).

CSS		FORMATO DI FILE
Nome completo	Cascaded Style Sheet	
Estensione/i	.css	
Magic number	–	
Tipo MIME	text/css	
Sviluppato da	World Wide Web Consortium	
Tipologia di standard	aperto, estendibile, *de iure*, testuale	
Livello metadati	2	
Derivato da	–	
Revisione	3	
Riferimenti	• W3C Recommendation *CSS 2.1 Specification*, 2011 • W3C Recommendation *CSS Basic User Interface module level 3 (CSS3 UI)*, 2018 • W3C Recommendation *CSS Color module level 3*, 2018 • W3C Recommendation *CSS Media Queries*, 2012 • validator.w3.org	
Conservazione	Sì; cfr. §2.8	
Racc. per la lettura	Generale; raccomandato per visualizzare pagine web	
Racc. per la scrittura	Generale; raccomandato per la presentazione di documenti sotto format di pagine web	

11. Il formato CSS effettua trasformazioni stilistiche di un documento HTML per permettere al suo contenuto di essere visualizzato graficamente, tramite un *web browser*, inclusi gli adattamenti "dinamici" del contenuto stesso. Una pagina web viene così rappresentata da vari file, tra i quali spiccano quelli delle seguenti due tipologie:

- il contenuto vero e proprio della pagina (testi, ipertesti e altri tipi di riferimenti ad altre pagine o contenuti), formato in HTML;

- la "presentazione" dei suddetti contenuti, formata in CSS.

12. Possono esistere più documenti CSS che adattano il medesimo contenuto a dispositivi di visualizzazione (p.es. stampa, monitor a bassa risoluzione, monitor ad alta risoluzione, dispositivi touch-screen, dispositivi per contenuti accessibili di vario tipo, …), ovvero adattano più contenuti diversi uniformandone lo stile. La versione di linguaggio CSS da usare dipende dalla versione HTML del contenuto: si raccomanda CSS3 per HTML5, ovvero CSS2 per XHTML.

MARKDOWN	FORMATO DI FILE
Nome completo	Markdown
Estensione/i	.md
Magic number	–
Tipo MIME	text/markdown
Sviluppato da	John Gruber, Aaron Swarz
Tipologia di standard	aperto, estendibile, *de iure*, testuale
Livello metadati	2
Derivato da	–
Revisione	0.28 (2017)
Riferimenti	• RFC-7763, RFC-7764 • spec.commonmark.org
Conservazione	Sì, se conservato insieme agli oggetti da esso riferiti
Racc. per la lettura	Generale; raccomandato per produzione di testi e ipertesti pubblicati online
Racc. per la scrittura	Generale; raccomandato per produzione di testi e ipertesti pubblicati online

13. *Markdown* ("md") è un linguaggio di markup pensato per scrivere contenuti testuali insieme ad una limitata quantità di ipertesti e di capacità "presentazionali". La sintassi semplificata del markdown è pensata per rendere il contenuto di un ipertesto accessibile e traducibile, in automatico, in linguaggi e sintassi che ne permettono un'adeguata presentazione e trasmissione, come ad esempio HTML, EPUB, OOXML, PDF o altro. Nella previsione che sempre più servizi delle PP.AA. saranno prodotti e accessibili online, il markdown si configura dunque come il linguaggio d'elezione per l'archiviazione a breve e lungo termine di questi contenuti, che sono perciò resi indipendenti dalla pagina web o dal file ove possano temporaneamente essere rappresentati. Gli ipertesti in markdown si prestano ad essere conservati, nella loro sintassi originale, sia all'interno di basi di dati generiche (cfr. §2.3) che in documenti informatici indipendenti, sottoforma di file con estensione .md.

MATHML	FORMATO DI FILE
Nome completo	MathML
Estensione/i	.mml, .xml
Specializzazione di	XML
Tipo MIME	text/mathml, text/mathml-renderer[*]
Sviluppato da	World Wide Web Consortium
Tipologia di standard	aperto, estendibile, *de iure*, testuale
Livello metadati	4
Derivato da	XML
Revisione	3.0 (2014)
Riferimenti	• ISO/IEC 40314:2016, *MathML version 3.0, 2nd Ed.* • W3C Recommendation *MathML version 3.0, 2nd Ed.*, 2014 • W3C *Math Home*
Conservazione	Sì; cfr. §2.8
Racc. per la lettura	Speciale; obbligatorio per testi tecnico-scientifici basati su XML
Racc. per la scrittura	Speciale; fortemente raccomandato per testi tecnico-scientifici; obbligatorio per quelli basati su XML

14. MathML è un dialetto di XML adatto alla rappresentazione di formule matematiche generiche; tale estensione di XML viene dunque utilizzata per testi scientifici in qualsivoglia documento informatico basato su XML, suoi dialetti (e.g. XHTML) o linguaggi "imparentati" (e.g. HTML).

2.2.1 Raccomandazioni per la produzione di documenti

1. Si raccomanda di usare per gli ipertesti 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 sitel 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.

2.3 Dati strutturati

1. In questa sezione si descrivono brevemente alcuni formati dedicati al trasporto di dati strutturati, intendendo con questa accezione riferirsi a formati ove la tipologica di contenuto non è predeterminata a propri. Esempi di applicazioni che fanno uso di dati strutturati sono le basi di dati (per le quali rappresentno qui i formati SQL e quelli relativi all'applicativo Microsoft® *Access*®). Si sottolinea che l'adeguatezza di una base di dati alla normativa vigente in materia di protezione dei dati personali (p.es. pseudonimìa) e privacy può essere indipendente dal formato di file adottato, mentre è fortemente caratterizzata dai criteri architetturali adottati durante la fase progettuale.

SQL		FORMATO DI FILE
Nome completo	Structured Query Language	
Estensione/i	.sql	
Specializzazione di	Datalog query language	
Tipo MIME	application/sql	
Sviluppato da	International Organization for Standardization	
Tipologia di standard	aperto, estendibile, *de iure*, testuale	
Livello metadati	4	
Derivato da	Datalog	
Revisione	SQL:2016	
Riferimenti	• RFC-6922 Famiglia di standard **9075** della ISO/IEC: • ISO/IEC 9075-1:2016, *SQL Part 1 – framework*	
Conservazione	Sì	
Racc. per la lettura	Speciale; raccomandato per basi di dati relazionali	
Racc. per la scrittura	Speciale; raccomandato per basi di dati relazionali	

2. I file SQL servono a contenere configurazioni e tabelle per basi di dati relazionali, complete o parziali, sotto forma del loro linguaggio di programmazione comune. Ogni file SQL descrive la formazione della base di dati, "da zero" o a partire da una base supposta già esistente: fornendo un tale file ad un gestore di basi di dati vengono perciò costituite le sue tabelle. Vice versa, una o più tabelle possono essere archiviate effettuandone uno "scarico" (*dump* in inglese) in un file SQL che descrive come il contenuto dello scarico può essere formato, da zero, in un nuovo gestore di basi di dati che interpreti il medesimo linguaggio. A tale scopo bisogna dunque specificare che SQL è in realtà un ceppo linguistico, da cui sono derivati molteplici dialetti di SQL, differenziati a seconda dell'applicativo –commerciale o meno– che funge da gestore di basi di dati. Si raccomanda quindi, per l'archiviazione a lungo termine e l'interscambio, di utilizzare sempre il formato SQL standardizzato dalla ISO (e riportato in tabella).
3. Ove non possibile, va sempre indicata la versione esatta del linguaggio SQL adottato, incluso preferenzialmente il nome e la versione completa dell'applicativo di gestione (MySQL, Microsoft® SQL™, ecc.).

Access® 2007		Formato di file
Nome completo	Microsoft® *Access* ® Connectivity Engine	
Estensione/i	.accdb	
Magic number	0x00010000 Standard ACE DB	
Tipo MIME	application/msaccess	
Sviluppato da	Microsoft Corporation	
Tipologia di standard	proprietario (chiuso), *de facto*, binario	
Livello metadati	3	
Derivato da	MS-MDB	
Revisione	2017	
Riferimenti	• Microsoft, *Which* Access® *file format should I use? – the .accdb file format* (2019)	
Conservazione	No	
Racc. per la lettura	Speciale; raccomandato per piccole basi di dati	
Racc. per la scrittura	Speciale; sconsigliato	

4. Il formato proprietario usato da Microsoft per il suo gestore di basi di dati *Access*®, a partire dalla versione 2007, permette di archiviare non soltanto tabelle, ma anche righe di codice SQL e dati nonstrutturati. Tuttavia, esseno il formato proprietario e a sorgente chiusa, viene elencato qui in caso vi siano organizzazioni con basi di dati basati sugli applicativi Microsoft, affinché migrino i dati un un formato non proprietario ovvero, qualora non altrimenti possibile, dal vecchio formato .mdb (si legga più avanti) a quello attuale .accdb.

MS-MDB	FORMATO DI FILE
Nome completo	Microsoft® *Access*® Binary file format
Estensione/i	.mdb
Magic number	0x00010000 Standard Jet DB
Tipo MIME	application/msaccess
Sviluppato da	Microsoft Corporation
Tipologia di standard	proprietario (chiuso), *de facto*, deprecato, binario
Livello metadati	3
Derivato da	—
Revisione	2017
Riferimenti	• Microsoft, *Which* Access® *file format should I use? – the .mdb file format* (2019)
Conservazione	No
Racc. per la lettura	Generico; valutare riversamento in formato aperto
Racc. per la scrittura	Sconsigliato; valutare un riversamento in altro formato

5. Il formato proprietario MDB è stato usato in precedenza da Microsoft per il suo gestore di basi di dati *Access*,® fino alla versione 2003. È ora un formato deprecato.

ODB	FORMATO CONTENITORE
Nome completo	Open Document Format for Database
Estensione/i	.odb
Specializzazione di	XML imbustato dentro ZIP
Tipo MIME	application/vnd.oasis.opendocument.[data]base
Sviluppato da	Organization for the Advancement of Structured Information Standards
Tipologia di standard	aperto, estendibile, *de iure*, binario, deprecato
Livello metadati	3
Derivato da	—
Revisione	1.2 (2015)
Riferimenti	Famiglia di standard **26300** della ISO/IEC. • OASIS, *Open Document Format for Office Applications (OpenDocument)*, v1.2 (2015)
Conservazione	No
Racc. per la lettura	Generico; raccomandato solo a scopo di riversamento
Racc. per la scrittura	Sconsigliato; valutare riversamento in altro formato

6. Specializzazione del formato OpenDocument per la rappresentazione di basi di dati, ODP è un formato deprecato e va perciò evitata la produzione di nuovi file;

inoltre, si consiglia di valutare il riversamento di basi di dati preesistenti in questo formato.

JSON		FORMATO DI FILE
Nome completo	JavaScript Object Notation	
Estensione/i	.json	
Specializzazione di	JavaScript	
Tipo MIME	application/json	
Sviluppato da	dominio pubblico	
Tipologia di standard	aperto, estendibile, *de iure*, testuale	
Livello metadati	4	
Derivato da	JavaScript	
Revisione	2018	
Riferimenti	• RFC-8259 • ECMA-404 • json.org • ISO/IEC 8825-8:2018, *ASN.1 encoding rules, part 8: JSON*	
Conservazione	Sì, se adottato insieme a uno schema JSON	
Racc. per la lettura	Generico; obbligatorio; eventualmente subordinato a ulteriori obblighi di conformità con determinati schemi.	
Racc. per la scrittura	Fortemente raccomandato per documenti contenenti dati sia strutturati che non strutturati; eventualmente subordinato a ulteriori obblighi di conformità con determinati schemi.	

7. Il formato *JavaScript Object Notation* (JSON) è un altro formato estremamente versatile ed estendibile che, come XML, è usato in moltissimi ambiti informatici e viene dunque impiegato per svariate applicazioni. Come definito nello standard RFC-8259 del 2018, in un file JSON le evidenze informatiche sono rappresentate gerarchicamente in più livelli:

- Ogni livello (compresa la "radice") è delimitato da parentesi graffe '{' e '}'.
- In ogni livello, ogni elemento (tranne l'ultimo) è separato dal successivo da una virgola ',.'.
- Ogni elemento è costituito da una coppia nome–valore, separati fra loro da due punti ':', ove il nome è sempre una stringa (delimitata da virgolette '"') e il valore può essere:
    - anch'esso una stringa (delimitata da virgolette '"');
    - un numero intero o in virgola mobile[13];

[13] Sono ammessi segni '+' e '-', punto decimale '.' e un numero arbitrario di cifre, ma non notazioni scientifiche (e.g. quella esponenziale).

- un booleano rappresentato dalle parole chiavi ‘true’ ovvero ‘false’;
- un *array* di valori, delimitato da parentesi quadre ‘[’ e ‘]’, separati (tranne l’ultimo) da virgole ‘,’ e con ciascun valore essendo di un qualsiasi tipo in questo elenco,
- un oggetto JSON (cioè sottolivello di questa medesima struttura);
- un tipo di dato che non indica nulla, rappresentato dalla parola chiave ‘null’.

- Tra i caratteri di delimitazione[14] sopra elencati, possono essere aggiunti un qualsivoglia numero di caratteri di spaziatura e interruzione di linea (soprattutto nella rappresentazione di tale struttura dati in un file), ottenendo un’equivalenza del medesimo contenuto.
- L’ordine degli elementi al medesimo livello di un *array* o di una struttura JSON è indifferente.

8. Un esempio di documento informatico rappresentato in formato JSON è dato dal seguente elenco di attributi di identificazione elettronica:

```
{
    "id":"cae8877c-e533-4d8a-9f38-d7f219392b1f",
    "firstName":"Mario",
    "lastName":"Rossi",
    "fiscalCode":"RSSMRA75L01H501A",
    "phoneNum":["+393201234567","+393398901234"],
    "children":[
        "id":"cae8877c-e533-4d8a-9f38-d7f219392b1f",
        "firstName":"Luigi",
        "lastName":"Bianchi",
        "fiscalCode":"RSSMRA75L01H501A",
        "phoneNum":["+393332468013"],
        "children":[
            ............
        ],
        "alive?":true
    ],
    "alive?":no
}
```

9. Un vantaggio di JSON rispetto a XML è la semplicità della sintassi che, pur non offrendo capacità avanzate quali i *namespace* o la possibilità di “attraversare” la struttura ad albero (offerte da XML), dispone di una struttura più semplice da interpretare –soprattutto da parte di processi automatici– senza perdere la caratteristica fondamentale della leggibilità da parte dell’uomo. Un impiego particolare di JSON è in alcuni tipi di basi di dati che necessitano uno scambio di tabelle potenzialmente grandi e complesse ma non necessariamente relazionate fra

[14] Indicati, in questo elenco e come al solito nell’Allegato, mediante caratteri a spaziatura fissa.

loro: in tali casi viene solitamente considerata una rappresentazione di interscambio mediante JSON piuttosto che i formati nativi dei altri tipi di basi di dati (p.es. SQL). Un'ulteriore caratteristica del JSON nel rappresentare basi di dati è che ogni dato può avere una struttura completamente diversa dagli altri; per questo motivo tale formato è ideale per l'archiviazione l'elaborazione, in automatico, di dati cosiddetti "non strutturati".

JSON-LD	FORMATO DI FILE
Nome completo	JavaScript Object Notation for Linked Data
Estensione/i	.jsonld
Specializzazione di	JSON
Tipo MIME	application/ld+json
Sviluppato da	dominio pubblico
Tipologia di standard	aperto, estendibile, *de iure*, testuale
Livello metadati	2
Derivato da	JavaScript
Revisione	2014
Riferimenti	• W3C Recommendation *JSON-LD 1.0*, 2014 • W3C Recommendation *JSON-LD 1.0 Processing Algorithms and API*, 2014
Conservazione	Sì
Racc. per la lettura	Generico; obbligatorio; eventualmente subordinato ad obblighi di conformità a schemi.
Racc. per la scrittura	Fortemente raccomandato per documenti contenenti Open Data generati con procedure automatizzate

10. Un particolare dialetto di JSON è costituito da JSON-LD, che segue una sintassi più rigida organizzata per poter gestire dati strutturati e collegati fra loro da relazioni astratte il più generiche possibili. Nell'ambito delle PP.AA., ad esempio, l'utilizzo di questo formato è –al pari di altri contenuti in questo Allegato– fortemente consigliato per la generazione di Open Data, particolarmente attraverso procedure automatizzate.

CSV	FORMATO DI FILE
Nome completo	Comma-Separated Value
Estensione/i	.csv
Magic number	–
Tipo MIME	text/csv
Sviluppato da	dominio pubblico
Tipologia di standard	aperto, estendibile, *de iure*, testuale
Livello metadati	1
Derivato da	tabelle preformattate in FORTRAN77

Revisione	2005
Riferimenti	• RFC-4180 • RFC-7111 • W3C Recommendation *Model for tabular data and metadata on the Web*, 2015 • dati.gov.it
Conservazione	Sì
Racc. per la lettura	Generale; obbligatorio
Racc. per la scrittura	Raccomandato per documenti contenenti dati strutturati leggibili dall'uomo (inclusi gli Open Data)

11. Uno dei formati più semplici per la rappresentazione di dati *fortemente* strutturati è il CSV che è un file testuale ove i dati sono rappresentati in una tabella:

- Ogni riga della tabella è una linea del file, delimitata da un'opportuna evidenza di fine linea.
- Le colonne, all'interno di ogni linea, si distinguono perché delimitate da caratteri specifici (ad esempio la virgola ',' che da il nome al formato, ma sono possibili anche altre combinazioni di caratteri di interpunzione o di spaziatura).
- Opzionalmente, la prima linea del file (indicata con un carattere iniziale particolare e con la medesima distinzione in colonne del resto delle linee) costituisce una legenda circa il significato dei valori delle colonne.

12. Il formato CSV è uno dei formati d'elezione per la rappresentazione degli Open Data, il cui archivio nazionale per i dati aperti pubblici è dati.gov.it.

JWT		CODEC
Nome completo	JSON Web Token	
Estensione/i	—	
Specializzazione di	JSON	
Tipo MIME	application/jwt	
Sviluppato da	pubblico dominio	
Tipologia di standard	aperto, estendibile, *de iure*, binario	
Derivato da	JSON	
Revisione	2016	
Riferimenti	• RFC-7797 • jwt.io • ISO/IEC 29500-6:2017 • RFC-7515 • RFC-7516	
Conservazione	Sì, se da JSON-LD o conservato insieme a schema	
Racc. per la lettura	Speciale; consigliato per flussi digitali in richieste POST	

Racc. per la scrittura	Speciale; consigliato per flussi digitali in richieste POST

13. Il *JSON Web Token* (JWT), definito dall'RFC-7797, è un codec[15] utilizzato per documenti informatici già organizzati in una struttura JSON; tale codifica ottimizza la dimensione dell'evidenza informatica e consente di trasferirla come un flusso digitale (cfr. §1.1.1) in tempo reale o attraverso un canale che permette un insieme limitato di caratteri — ad esempio inserendo l'evidenza in richieste HTTP di tipo POST. Il JWT (analogamente al JSON-LD) richiede che nella struttura JSON testuale possano essere presenti alcuni elementi, facoltativi ma dal significato predefinito, che semplificano il trasporto e la decodifica dell'evidenza JWT stessa.
Lo standard JWT include un sistema di controllo della propria integrità e può, facoltativamente, essere accompagnato da cifratura (per la confidenzialità), da apposizione di firma o sigillo elettronico nel pacchetto, o da entrambe. Le evidenze informatiche JWT firmate sono, impropriamente, chiamate *JSON Web Signature* (JWS) e definite nell'RFC-7515; le evidenze informatiche JWT cifrate sono, impropriamente, chiamate *JSON Web Encryption* (JWE) e definite nell'RFC-7516.

OOXML	FORMATO CONTENITORE
Nome completo	Office Open XML
Estensione/i	.docx, .xlsx, .pptx, …
Specializzazione di	XML imbustato dentro ZIP
Tipo MIME	–
Sviluppato da	Microsoft Corporation
Tipologia di standard	aperto, estendibile, *de iure*, testuale
Livello metadati	4; cfr. §2.8
Derivato da	OpenXML; Microsoft® COM Structured Storage
Revisione	12.0 (2019)
Riferimenti	• ISO/IEC 29500-1:2016, *fundamentals and markup language reference* • ISO/IEC 29500-2:2012, *open packaging conventions* • ISO/IEC 29500-3:2015, *markup compatibility and extensibility* • ISO/IEC 29500-4:2016, *transitional migration features* • ECMA-376: *Office Open XML File Formats*, 5th ed., 2016 • officeopenxml.com
Conservazione	Sì, solo profilo Strict; cfr. §2.8
Racc. per la lettura	–

[15] Per la precisione, la codifica binaria è ottenuta mediante algoritmo *Base64*, specificato in RFC-4648.

Racc. per la scrittura –

14. Vengono elencati due formati contenitore utilizzati dalle principali suite di applicativi per ufficio: Microsoft® *Office*® e *LibreOffice*. Tali buste contengono in realtà un pacchetto di file che rappresenta, complessivamente, un documento informatico (testo impaginato, foglio di calcolo, presentazione, basi di dati, contenuto multimediale, o altro) mediante più file strutturati in un filesystem virtuale, con gerarchie di cartelle predeterminate. Il pacchetto, una volta formato in memoria, viene poi compresso in un unico file mediate algoritmi noti (tra quelli elencati in §2.13), o loro varianti. I file contenuti nel pacchetto descrivono sia i metadati interni del documento, che parte del documento informatico stesso: ad esempio testi, collegamenti ipertestuale e la loro rappresentazione grafica (impaginazione, colore, tipografia, ecc.) e sono tipicamente in formato XML. Eventuali contenuti ipertestuali (pagine web, immagini, suono, video, contenitori crittografici ecc.) sono rappresentati da file in altri formati –tipicamente aperti o proprietari ma liberi– (HTML, PNG, WAV, PEM, ecc.) a seconda della necessità. Il pacchetto di file viene di solito formato e archiviato nello storage soltanto previa compressione in un unico file, mentre al caricamento da parte di un applicativo viene scompattato in memoria. La strutturazione in un pacchetto di file consente una migliore gestione del documento particolarmente nei casi di dimensioni elevate del file: una lieve modifica parziale di un documento informatico molto grande comporta la modifica di un sottoinsieme di file costituenti il pacchetto, non di tutti, perciò lo sforzo computazionale si riduce, così come i rischi di corruzione dell'intero file in caso di problemi durante la scrittura di file così grandi.

15. *Office Open XML*[16] (*OOXML*) è il contenitore generico utilizzato prevalentemente dalle versioni più recenti (dalla 2007 in poi) della suite applicativa Microsoft® *Office*,® che si specializza a seconda della tipologia di documento da contenere (e di applicativo della suite). Un documento in formato OOXML è in realtà costituito da un pacchetto di file che sono poi compressi in un'unica busta ZIP (cfr. §2.13), rinominata con estensione differente a seconda della specializzazione.
16. In questo Allegato sono descritte tre specializzazioni di OOXML ai documenti impaginati (.docx, §2.1), ai fogli di calcolo e alle presentazioni multimediali (rispettivamente .xlsx e .pptx, §2.5)

[16] Da non confondersi con *OpenOffice.org XML* — formato simile utilizzato da versioni obsolete dell'omonima suite applicativa.

OpenDocument	Formato contenitore
Nome completo	Open Document Format for Office Applications
Estensione/i	.odt, .ods, .odp, .odg, .odi, .odf
Specializzazione di	XML imbustato dentro ZIP
Tipo MIME	—
Sviluppato da	Organization for the Advancement of Structured Information Standards
Tipologia di standard	aperto, estendibile, *de iure*, binario
Livello metadati	4
Derivato da	—
Revisione	1.2 (2015)
Riferimenti	Famiglia di standard **26300** della ISO/IEC: • ISO/IEC 26300-1:2015, *ODF for Office Applications v1.2 – Part 1: OpenDocument Schema* • ISO/IEC 26300-3:2015, *ODF for Office Applications v1.2 – Part 3: Packages* • OASIS, *Open Document Format for Office Applications (OpenDocument)*, v1.2 (2015)
Conservazione	Sì; cfr. §2.8
Racc. per la lettura	—
Racc. per la scrittura	—

17. OOXML è dotato di vari "profili", che possono essere più o meno interoperabili. Per tutti i documenti della P.A., si consiglia perciò il profilo Strict, che è più restrittivo, ma consente di eliminare alcune estensioni "proprietarie" che possono ridurre l'interoperabilità dei formati.

18. Come per OOXML, anche OpenDocument è un contenitore generico di quelli sopraelencati, per archiviare documenti prodotti da suite applicative open source — prima fra tutte *LibreOffice*.

In questo allegato sono descritte alcune specializzazioni di OpenDocument ai documenti impaginati (.odt, §2.1), ai fogli di calcolo e alle presentazioni multimediali (rispettivamente .ods e .odp, §2.5), ai

CFB	Formato contenitore
Nome completo	Compound File Binary file format
Estensione/i	.doc, .xls, .ppt, .pst; .aaf, …
Magic number	0xD0CF11E0A1B11AE1
Tipo MIME	—
Sviluppato da	Microsoft Corporation
Tipologia di standard	proprietario, estendibile, *de facto*, binario, deprecato

Livello metadati	–
Derivato da	Microsoft® COM Structured Storage
Revisione	9.0 (2018)
Riferimenti	• Microsoft, *[MSD-CFB]: Compound File Binary file format* v9.0 (2018)
Conservazione	No; cfr. §2.8
Racc. per la lettura	–
Racc. per la scrittura	–

19. CFB è invece il contenitore di file su cui Microsoft® si è basata per le versioni obsolete della suite *Office*® (cfr. § 2.1 e §2.5). È anche il formato contenitore da cui deriva AAF (cfr. §2.12). Il formato CFB rappresenta all'interno un pacchetto di file organizzati in un filesystem virtuale vagamento ispirato al FAT[17] (anziché in un archivio compresso come fanno altri formati). Una delle principali problematiche del CFB, soprattutto nel caso di documenti informatici di grandi dimensioni, è ereditata proprio dal FAT, essendo caratterizzato dall'estrema facilità con file contenuti del pacchetto soffrono di frammentazione interna. Ad un livello più superficiale, ogni applicativo della suite Microsoft® *Office*® versione '2003' o antecedenti utilizza una propria serie ti tipi di file, dotati di estensioni differenti (p.es. .doc, .xls, .ppt, …, cfr. §1.1.2), ma tutti basati su CFB;[18] a partire dalla versione '2007' la suite ha abbandonato i formati basati su CFB in favore di quelli basati su OOXML (Office Open XML), producendo nuovi formati, stavolta strutturalmente differenti ma gemellati ai precedenti per via delle estensioni di file, ottenute quasi sempre aggiungendo una semplice 'x': .docx, .xlsx, .pptx, ….

2.3.1 Raccomandazioni per la produzione di documenti

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

[17] Acronimo di *File Allocation Table*, dal nome dell'evidenza informatica, contenuta (o replicata più volte) in dispositivo di storage a blocchi inizializzato con tale filesystem, per indicizzarne tutti i file contenuti e i loro metadati esterni (cfr. §1.1.2). Esistono diversi filesystem basati su FAT — alcuni aperti (FAT16, FAT32), altri coperti da licenze d'uso di proprietà di Microsoft Corporation (exFAT). Per ulteriori dettagli consultare it.wikipedia.org/wiki/File_Allocation_Table e i collegamenti ipertestuali ivi contenuti.

[18] A ciascuno di tali formati si affiancano delle varianti, strutturalmente identiche, per rappresentare specializzazioni dei medesimi documenti, quali ad esempio i "modelli di documento" (*template*, in inglese), per i quali cambia solo l'estensione del file (tipicamente una 't' in sostituzione dell'ultimo carattere: p.es. .dot e .xlt).

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 transito" (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:
    - dati personali e privacy: cfr. Regolamento (UE) Nº 679/2016 ("GDPR") del Parlamento europeo e del Consiglio e il D.Lgs. 101/2018;
    - 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.

2. Si raccomanda perciò alle PP.AA. di effettuare un'adeguata valutazione di interoperabilità, che tenga in considerazione anche dei sopracitati aspetti.

2.4 Posta elettronica

EML	FORMATO DI FILE
Nome completo	Electronic Mail Format
Estensione/i	.eml
Magic number	–
Tipo MIME	application/email
Sviluppato da	comunità open source
Tipologia di standard	aperto, estendibile, *de facto*, testuale
Livello metadati	1
Derivato da	RFC-822
Revisione	2008
Riferimenti	• RFC-5322 • RFC-2822 • US Library of Congress, *Email (EMF)* (2014)
Conservazione	Sì; cfr. §2.8

Racc. per la lettura	Generico; obbligatorio per singoli messaggi email
Racc. per la scrittura	Generico; obbligato per singoli messaggi email

1. Il formato EML rappresenta interamente l'evidenza informatica costituente un singolo messaggio di posta elettronica MIME, così come definito negli standard RFC-5322 e RFC-2822.

MBOX	FORMATO CONTENITORE
Nome completo	"default" *mbox* database format
Estensione/i	.mbox
Specializzazione di	EML
Tipo MIME	application/mbox
Sviluppato da	comunità open source
Tipologia di standard	aperto, estendibile, *de facto*, testuale
Livello metadati	3
Derivato da	sistema operativo UNIX
Revisione	2005
Riferimenti	• RFC-4155 • RFC-2822 • en.wikipedia.org/wiki/Mbox
Conservazione	Sì; cfr. §2.8
Racc. per la lettura	Generico; raccomandato per caselle di messaggi email
Racc. per la scrittura	Generico; raccomandato per caselle di messaggi email

2. Così come EML rappresenta, integralmente un singolo messaggio di posta elettronica, il formato MBOX può usarsi per contenervi diversi messaggi di posta elettronica, organizzati su più livelli.

MS-PST	FORMATO DI FILE
Nome completo	Microsoft® *Outlook*® Personal Folder file
Estensione/i	.pst
Magic number	0xD0CF11E0A1B11AE1
Tipo MIME	application/vnd.ms-outlook
Sviluppato da	Microsoft Corporation
Tipologia di standard	proprietario (libero), estendibile, *de facto*, binario
Livello metadati	3
Derivato da	Microsoft® Compound File binary format
Revisione	7.0 (2018)
Riferimenti	• Microsoft, *[MS-PST]: Outlook Personal Folders (.pst) file format* v7.0 (2018)
Conservazione	No; cfr. §2.8

Racc. per la lettura	Generico; raccomandato per importare rubriche, ecc.
Racc. per la scrittura	Specifico; raccomandata migrazione dell'applicativo e, successivamente, dell'intero formato

3. Il formato proprietario usato in Microsoft® *Outlook*® (applicativo di rubrica messaggi e contatti) è un'alternativa al formato MBOX per memorizzare non soltanto intere caselle di posta (a loro volta strutturabili in sottocartelle), ma anche per memorizzare altre tipologie di dati utilizzabili in *Outlook*®, quali ad esempio schedari dei contatti, note personali, ecc. Il formato PST è anch'esso costituito da un pacchetto di file compressi in un unico file.

2.4.1 Raccomandazioni per la produzione di documenti

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.

2.5 Fogli di calcolo e presentazioni multimediali

1. Per formati di file maggiormente utilizzati per gli applicativi integrativi "da ufficio" si rimanda alle considerazioni già fatte in § 2.1 relativamente agli applicativi di videoscrittura ad essi affini, in particolare quelle al capoverso 10 del paragrafo in merito all'utilizzo delle fonti tipografiche.

EXCEL® 2007	FORMATO DI FILE
Nome completo	SpreadsheetML OOXML Extension
Estensione/i	.xlsx, .xltx
Specializzazione di	XML imbustato dentro ZIP
Tipo MIME	application/vnd.openxmlformats-officedocument.spreadsheetml.sheet
Sviluppato da	Microsoft Corporation
Tipologia di standard	proprietario (libero), estendibile, *de facto*
Livello metadati	3
Derivato da	Office Open XML; Microsoft® *Excel* ®
Revisione	16.0 (2018)

Riferimenti	• Microsoft, *Excel (.xlsx) extensions to SpreadsheetML file format* v16.0 (2018) • officeopenxml.com, *Anatomy of a SpreadsheetML file*
Conservazione	Sì, solo profilo Strict; cfr. §2.8
Racc. per la lettura	Generico con riconoscimento obbligatorio
Racc. per la scrittura	Raccomandato nel profilo Strict

2. SpreadsheetML è il dialetto XML usato nei file di metadati all'interno di un pacchetto compresso in formato OOXML (§2.3), specializzato per la rappresentazione di fogli di calcolo (estensione .xlsx). È stato introdotto con la versione 2007 di Microsoft® *Office*®, ma è compatibile con moltissimi altri applicativi. L'unico profilo raccomandato di OOXML per SpreadsheetML è Strict.
3. Si segnala che Microsoft® *Excel*® presenta un *glitch* (i.e. un *bug* volutamente introdotto) considerando erroneamente l'anno 1900 d.C. come bisestile; il formato SpreadsheetML (profilo Transitional), di per sé, non corregge questo comportamento, permettendo il salvataggio, nelle celle di un foglio di calcolo, della data inesistente del 29 febbraio 1900. Tale inesattezza è però imputabile ad *Excel*® e non inficia la capacità di altri applicativi che fanno uso di codesto formato di file (p.es. Google *Documents*) di adeguarsi ad un corretto calcolo degli anni bisestili (correggendo automaticamente le date inesatte nei fogli di calcolo).

POWERPOINT® 2007		FORMATO DI FILE
Nome completo	PresentationML OOXML Extension	
Estensione/i	.pptx, .ppsx, .potx	
Specializzazione di	XML imbustato dentro ZIP	
Tipo MIME	application/vnd.openxmlformats-officedocument.presentationml.presentation	
Sviluppato da	Microsoft Corporation	
Tipologia di standard	proprietario (libero), estendibile, *de facto*	
Livello metadati	3	
Derivato da	Office Open XML; Microsoft® *PowerPoint* ®	
Revisione	15.0 (2018)	
Riferimenti	• Microsoft, *PowerPoint (.pptx) extensions to OOXML file format* v15.0 (2018) • officeopenxml.com, *Anatomy of a PresentationML file*	
Conservazione	Sì, solo profilo Strict; cfr. §2.8	
Racc. per la lettura	Generico con riconoscimento obbligatorio	
Racc. per la scrittura	Raccomandato nel profilo Strict	

3. PresentationML è il dialetto XML usato nei file di metadati all'interno di un pacchetto compresso in formato OOXML (§2.3), specializzato per la rappresentazione di presentazioni multimediali (estensione .pptx). È stato introdotto con la versione 2007 di Microsoft® *Office*®, ma è compatibile con moltissimi altri applicativi. L'unico profilo raccomandato di OOXML per PresentationML è Strict.

MS-XLS	FORMATO DI FILE
Nome completo	Microsoft® *Excel* ® Binary file format
Estensione/i	.xls
Magic number	0xD0CF11E0A1B11AE1
Tipo MIME	application/vnd.ms-excel
Sviluppato da	Microsoft Corporation
Tipologia di standard	proprietario (libero), estendibile, *de facto*, deprecato, binario
Livello metadati	3
Derivato da	Microsoft® Compound File binary format
Revisione	8.0 (2018)
Riferimenti	• Microsoft, *[MS-XLS]: Excel Binary file format (.xls) structure* v8.0 (2018)
Conservazione	No; cfr. §2.8
Racc. per la lettura	Obbligatorio con riversamento raccomandato
Racc. per la scrittura	Sconsigliato

5. Il foramto XLS è una specializzazione usata nei file di metadati all'interno di un pacchetto compresso in formato CFB (§2.3), specializzato per la rappresentazione di fogli di calcolo (estensione .xls). È stato utilizzato, come formato principale per tali documenti fino alla versione 2003 di Microsoft® *Office*®, ma è compatibile con moltissimi altri applicativi. Nonostante l'obbligo per le PP.AA. di accettare e aprire documenti in questo formato, si raccomanda di non formarne altri esemplari e di valutare il riversamento di fogli di calcolo preesistenti in altro formato della stessa tipologia: in quest'ordine, OpenDocument Spreadsheet (.ods) ovvero SpreadsheetML (.xlsx).

MS-PPT	FORMATO DI FILE
Nome completo	Microsoft® *PowerPoint* ® Binary format
Estensione/i	.ppt
Magic number	0xD0CF11E0A1B11AE1
Tipo MIME	application/vnd.ms-powerpoint

Sviluppato da	Microsoft Corporation
Tipologia di standard	proprietario (libero), estendibile, *de facto*, deprecato, binario
Livello metadati	3
Derivato da	Microsoft® Compound File binary format
Revisione	6.0 (2018)
Riferimenti	• Microsoft, *[MS-PPT]: PowerPoint (.ppt) Binary file format* v6.0 (2018)
Conservazione	No; cfr. §2.8
Racc. per la lettura	Obbligatorio con riversamento raccomandato
Racc. per la scrittura	Sconsigliato

6. Il foramto PPT è una specializzazione usata nei file di metadati all'interno di un pacchetto compresso in formato CFB (§2.3), specializzato per la rappresentazione di presentazioni multimediali (estensione .ppt). È stato utilizzato, come formato principale per tali documenti fino alla versione 2003 di Microsoft® *Office*®, ma è compatibile con moltissimi altri applicativi. Nonostante l'obbligo per le PP.AA. di accettare e aprire documenti in questo formato, si raccomanda di non formarne altri esemplari e di valutare il riversamento di presentazioni preesistenti in altro formato della stessa tipologia: in quest'ordine, OpenDocument Presentation (.odp) ovvero PresentationML (.pptx).

ODS	FORMATO CONTENITORE
Nome completo	Open Document Format for Office Spreadsheets
Estensione/i	.ods
Specializzazione di	XML imbustato dentro ZIP
Tipo MIME	—
Sviluppato da	Organization for the Advancement of Structured Information Standards
Tipologia di standard	aperto, estendibile, *de iure*, binario
Livello metadati	3
Derivato da	—
Revisione	1.2 (2015)
Riferimenti	Famiglia di standard **26300** della ISO/IEC. • OASIS, *Open Document Format for Office Applications (OpenDocument)*, v1.2 (2015)
Conservazione	Sì; cfr. §2.8
Racc. per la lettura	Generico con riconoscimento obbligatorio
Racc. per la scrittura	Fortemente raccomandato

7. Il formato OpenDocument Spreadsheet è una specializzazione dell'omonimo formato (§2.3) per rappresentare fogli di calcolo (estensione .ods). È attualmente

utilizzato dalla suite open source di applicativi da ufficio *LibreOffice*, anche se è pienamente utilizzabile in Microsoft® *Office*®, in OpenOffice.org e in altri applicativi che elaborano documenti di questo tipo.

ODP	FORMATO CONTENITORE
Nome completo	Open Document Format for Presentations
Estensione/i	.odp
Specializzazione di	XML imbustato dentro ZIP
Tipo MIME	–
Sviluppato da	Organization for the Advancement of Structured Information Standards
Tipologia di standard	aperto, estendibile, *de iure*, binario
Livello metadati	3
Derivato da	–
Revisione	1.2 (2015)
Riferimenti	Famiglia di standard **26300** della ISO/IEC. • OASIS, *Open Document Format for Office Applications (OpenDocument)*, v1.2 (2015)
Conservazione	Sì; cfr. §2.8
Racc. per la lettura	Generico con riconoscimento obbligatorio
Racc. per la scrittura	Fortemente raccomandato

8. Il foramto OpenDocument Presentation è una specializzazione dell'omonimo formato (§2.3) per rappresentare fogli di calcolo (estensione .odp). È attualmente utilizzato dalla suite open source di applicativi da ufficio *LibreOffice*, anche se è pienamente utilizzabile in Microsoft® *Office*,® in OpenOffice.org e in altri applicativi che elaborano documenti di questo tipo.

2.5.1 Raccomandazioni per la produzione di documenti

1. Si raccomanda alle PP.AA. 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.
2. 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,[19] purché tali formati siano

[19] Si veda a tale proposito le Circolari AGID N°2/2018 e N°3/2018.

interoperabili e disponibili su sistemi informativi senza vincoli o particolari requisiti tecnologi (ad esempio, totalmente fruibili attraverso browser web).

2.6 Immagini raster

1. Si dicono *raster* tutte le immagini che sono descritte come un insieme finito di punti (*pixel*) virtuali disposti regolarmente su una griglia rettangolare. Il numero di punti lungo i due lati del rettangolo (lunghezza e larghezza) costituiscono le dimensioni dell'immagine, mentre il numero di bit necessari a rappresentare le caratteristiche di ciascun pixel (espresso in bit/pixel) è chiamata profondità o (*bit-)depth* dell'immagine. Tali caratteristiche sono tipicamente le coordinate-colore (*code-values*) di ciascun pixel secondo un tipo di codifica matematica ove ogni possibile colore rappresentabile dalle immagini (*gamut*) è definito dalle coordinate di uno spazio-colore astratto, all'interno del quale la gamut dell'immagine rappresenta una varietà discreta.
2. A seconda delle esigenze di conservazione possono essere utili o necessari riferimenti che permettono di correlare la distribuzione spaziale dei pixel di un'immagine raster con dimensioni fisiche. Alcuni formati di file descrivono tali informazioni mediante metadati interni, che possono essere facoltativi o meno. La scelta del formato di file per usi professionali dipende spesso dalla possibilità di creare e, successivamente, fruire di tali metadati, rilevanti sia per raster che provengono da oggetti reali (come nel caso di fotografie, scansioni digitali e diagnostica medica), che per immagini di sintesi usate per la produzione oggetti reali (p.es. stampe o modelli di altro tipo).
3. In alcuni casi sono indicati le dimensioni equivalenti dell'immagine in unità di misura fisiche (lunghezza, larghezza, profondità; ovvero aree, volumi); in altri casi è indicata la risoluzione espressa come densità di punti per unità di misura lineare, quali punti/pollice ("dpi", cioè *dots per inch*) ovvero pixel/cm (nel caso dei raster i termini "punto" e "pixel" sono spesso intercambiabili).
4. Il numero di bit usati per rappresentare il colore dei pixel è chiamata profondità di colore (*colour-depth*) e coincide con l'intera profondità dell'immagine raster qualora la sola caratteristica rappresentata sia il colore dei pixel.
5. Altre volte sono rappresentate una o più di altre caratteristiche del pixel (in alternativa o in aggiunta al colore), quali:

- trasparenza (chiamata anche "alfa", *alpha* in inglese),
- densità ottica,
- riflettanza,
- intensità radiante,

- velocità angolare o lineare,

ciascuna di queste caratteristiche del pixel (inclusa ciascuna coordinata dello spazio-colore) sono chiamate, individualmente, "canali" dell'immagine; il numero di bit usati per rappresentare ciascun canale (qualora sia uguale per tutti i canali) permette di specificare la profondità dell'immagine, alternativamente, in bit/canale.

6. Ad esempio, un'immagine raster a colori che usa il tipico modello colorimetrico rosso-verde-blu e un canale alfa, tutti a 16 bit/canale (4 canali in tutto: RGBα) ha dunque una profondità complessiva di 64bit/pixel e una profondità-colore di 48 bit/pixel.

7. Molti formati di immagini raster possono –opzionalmente o obbligatoriamente– implementare algoritmi di compressione dati per ottenere una minore occupazione di spazio. Tali algoritmi sono *lossless* ovvero *lossy*, a seconda che la compressione sia reversibile (cioè sia possibile, decomprimendo, ritornare punto per punto all'immagine originale) o meno.

8. Un'altra possibilità per le immagini raster è una rappresentazione mediante sotto-campionamento, cioè utilizzando una codifica con un modello di spazio-colore a 3 componenti (anziché un modello RGB), ove il primo canale (luminanza) ha una risoluzione spaziale piena (cioè è codificato per ogni punto dell'immagine), mentre gli altri due canali, detti complessivamente di crominanza,[20] sono codificati con una minore risoluzione.

9. I formati di immagini raster maggiormente diffusi per usi non professionali, quali ad esempio web e fotografia amatoriale, sono: PNG, JPEG (.jpg, .jpeg), TIFF (.tif, .tiff), GIF.

10. I formati di immagini raster diffusi per vari usi professionali quali grafica, stampa industriale, fotografia e cinematografia, si differenziano dai precedenti per un ampio supporto –ed effettivo utilizzo– di metadati interni, prevalentemente di tipo colorimetrico. Tra questi formati si annoverano: OpenEXR (.exr), Adobe® DNG (.dng), JPEG2000 (.jp2k, .jp2c), DPX, Adobe® *Photoshop*® (.psd), ARRIRAW (.ari).

PNG		FORMATO DI FILE
Nome completo	Portable Network Graphics	
Estensione/i	.png	
Magic number	0x89 PNG 0x0D0A1A0A	
Tipo MIME	image/png	
Sviluppato da	ACME	
Tipologia di standard	aperto, *de iure*, binario, muto	
Livello metadati	3	

[20] Tali canali codificano le differenze cromatiche da un colore neutro –di solito qualche tipo di verde– la cui intensità è rappresentata dalla sola luminanza. Gli spazi-colore di questo tipo (p.es. $Y'c_Bc_R$ e $Y'uv$) sono nati per permettere la retrocompatibilità del segnale televisivo analogico a colori con quello in bianco e nero, affiancati da altre tecnologie indipendenti, quali il sotto-campionamento.

Derivato da	–
Revisione	1.2, 2ª edizione
Riferimenti	• ISO/IEC 15948:2004 • W3C Recommendation *PNG Specification (2nd Ed.)*, 2003 • RFC-2083 • www.libpng.org
Conservazione	Sì
Racc. per la lettura	Generico; con riconoscimento obbligatorio della v1.0
Racc. per la scrittura	Generico; fortemente raccomandato per immagini a 16 ovvero 48 bit/pixel (più eventuale trasparenza) senza particolari obblighi relativi ad altri metadati

11. Il formato PNG è particolarmente raccomandato per la rappresentazione di immagini raster che non hanno bisogno di essere accompagnate da metadati particolarmente complessi (come quelli colorimetrici, geometrico-fisico o ottici). La versione 1.0 di questo formato (cfr. RFC-2083) –ancora largamente utilizzata e, per questo motivo, ammessa parimenti per la produzione di nuovi documenti raster– supporta un insieme ristrettissimo di metadati, per la cui mancanza il formato di questa versione è poco usato in ambiti professionali. Sono invece supportate nativamente le profondità di 24 e 48 bit/pixel in tricromia, la trasparenza sotto forma di un ulteriore canale "alfa", nonché uno dei pochi formati non professionali ad ammettere la compressione dell'immagine mediante algoritmo *lossless*. La versione 1.2 del formato introduce alcuni metadati più professionali che permettono, ad esempio, di specificare autori e brevi descrizioni, orari di modifica, così come metadati colorimetrici (quali le coordinate di cromaticità delle primarie RGB e del punto di bianco, profili ICC integrati, dimensioni fisiche dell'immagine, ecc.).

JPEG	FORMATO DI FILE
Nome completo	JPEG File Interchange Format (JFIF)
Estensione/i	.jpg, .jpeg
Magic number	0xFFD8, 0xFFD8FFE00010 JFIF 0x0001
Tipo MIME	image/jpg, image/jpeg
Sviluppato da	Joint Photographic Experts Group
Tipologia di standard	aperto, estendibile, *de iure*, binario
Livello metadati	4
Derivato da	–
Revisione	2012
Riferimenti	• ITU-T Recommendation T.81, 1992 • ITU-T Recommendation T.871, 2011 • www.jpeg.org/jpeg • www.exif.org/Exif2-2.PDF

Conservazione	Sì, solo per immagini formate nativamente in JPEG
Racc. per la lettura	Generico con riconoscimento obbligatorio
Racc. per la scrittura	Generico; fortemente raccomandato per immagini fotografiche senza particolari vincoli qualitativi

12. Il formato JPEG è un altro "pilastro" storico per le immagini raster, permettendo una loro rappresentazione generica e, soprattutto, estremamente compatta per via dell'uso di un algoritmo di compressione *lossy*, in cui alcuni parametri algoritmici sono selezionabili in fase di formazione del file, permettendo così un bilanciamento tra "qualità" dell'immagine[21] e dimensione del file. Le immagini sono inoltre codificate sempre in uno spazio-colore del tipo *Y'uv* e sotto-campionate; non sono supportate le trasparenze. Nonostante queste carenze lo standard *EXIF* estende il formato per la sola rappresentazione di metadati interni, tra i quali possono essere presenti:

- informazioni sull'hardware o software usato per creare o modificare l'immagine (p.es. modello di fotocamera, ottiche, flash e altra apparecchiatura, con i loro parametri di scatto);
- data, ora ed eventuale posizione geografica della creazione;
- nome o altre caratteristiche dell'autore;
- eventuali licenze d'uso dell'immagine (Copyright©, Creative Commons, ecc.);
- informazioni colorimetriche (spazio-colore, punto di bianco, profilo ICC, ecc.);
- eventuale proiezione piana di uno spazio curvo (p.es. equi-rettangolare, altrimenti detta 'latitudine/longitudine' o, più impropriamente, "*360°*").

13. Qualora si disponga delle medesime immagini in un formato di maggiore qualità (in termini di risoluzione, di assenza di compressione *lossy*, o altro), si consiglia di non riversare mai in JPEG il medesimo contenuto (a meno che non si tratti di una seconda copia per altri scopi).

TIFF	FORMATO DI FILE
Nome completo	Tagged Image File Format
Estensione/i	.tiff, .tif
Magic number	II 0x2A00; MM 0x002A
Tipo MIME	image/tiff
Sviluppato da	Adobe Systems
Tipologia di standard	proprietario (libero), estendibile, *de iure*, binario

[21] Da intendersi qui come fedeltà dell'immagine compressa (con perdite dovute alla compressione) rispetto all'immagine originale non compressa. Tale quantità è misurata con opportune metriche, tra cui quelle basate sul rapporto segnale-rumore percepito, o pSNR.

Livello metadati	4
Derivato da	–
Revisione	2004
Riferimenti	• Adobe, *TIFF™ Revision 6.0*, 1992 • www.adobe.io/open/standards/TIFF Famiglia di standard **12234** della ISO: • ISO 12234-2:2001, TIFF/*EP* • ISO 12234-3:2016, *XMP* • ISO 12639:2004, TIFF/*IT* • RFC-2306, TIFF-*F* • RFC-3949, TIFF-*FX* • www.exif.org/Exif2-2.PDF
Compressione	Sì, senza compressione
Racc. per la lettura	Generico con riconoscimento obbligatorio
Racc. per la scrittura	Specifico per l'editoria; raccomandato per la produzione di immagini raster finali

13. Il formato TIFF è il capostipite dei formati raster professionali, pur essendo ancora considerato appartenente alla categoria dei formati generali per le immagini raster. Nato con poche varianti e metadati prevalentemente per rappresentare immagini pronte sia per l'invio telematico (tramite fax) che per la stampa professionale –monocromatica o "offset"– è stato successivamente esteso dalla Adobe, fino a diventare uno standard aperto e *de iure*, con moltissime caratteristiche aggiuntive:

- segmentazione dell'immagine nel file per linee ovvero per riquadri;[22]
- segmentazione a più livelli sovrapposti;[23]
- spazi-colore con un qualsivoglia numero di canali (inclusa la trasparenza, o canale *alfa*);
- metadati aggiuntivi circa la rappresentazione spaziale (unità di misura, ecc.);
- metadati aggiuntivi circa le condizioni ideali per la visualizzazione (caratteristiche del monitor) o per la stampa professionale (intento di rappresentazione, codici di particolari colori "spot", ecc.);
- compressione secondo vari algoritmi, generalmente *lossless*.

Infine, anche nelle immagini TIFF –come per le JPEG– possono essere inseriti metadati EXIF.

GIF		FORMATO DI FILE
Nome completo	Graphic Image file Format	

[22] Utili per ottimizzarne l'archiviazione su alcuni storage, la trasmissione attraverso specifici canali, la robustezza in caso di corruzione.
[23] Utile per immagini in fase di post-produzione, cioè quando sono semilavorati le cui modifiche devono essere reversibili.

Estensione/i	.gif
Magic number	GIF89a; GIF87a
Tipo MIME	image/gif
Sviluppato da	CompuServe
Tipologia di standard	deprecato, proprietario (libero), *de iure*, binario, muto
Livello metadati	1
Derivato da	–
Revisione	89a (1989)
Riferimenti	• W3C, *Graphics Interchange Format™*, ©1990 CompuServe
Conservazione	No
Racc. per la lettura	Generico con riconoscimento obbligatorio
Racc. per la scrittura	Sconsigliato, tranne che in due casi legati al web

14. Il formato GIF venne introdotto nella fine degli anni '80 per rappresentare immagini fotorealistiche quando le capacità dei computer di visualizzare simultaneamente molti colori erano limitate. Alcune delle novità tecniche di questo formato sono sopravvissute sino ad oggi, tanto che vi sono ancora molti casi d'uso soprattutto nelle pagine web. Anche se esistono oggi formati molto più adeguati per le immagini fotorealistiche, il formato, implementando fra le altre cose un ana compressione *lossless*, è ancora tecnicamente valido per immagini con le seguenti caratteristiche:

- numero colori fino a un massimo di 256, scelti da una tavolozza con 32 bit di profondità;[24]
- supporto per una trasparenza assente ovvero totale,[25] senza diversi gradi di opacità;
- supporto per più immagini che si alternano in sequenza (con cadenza preselezionata) per rappresentare uno sprite dotato di caratteristiche di animazione basilari (si parla, in questo caso di "GIF animate");[26]

Le immagini con bassissimi requisiti di animazione, trasparenza e numero di colori (tipicamente loghi o porzioni di elementi grafici minimalisti), come quelle che si trovano in diversi siti web, spesso salvate in questo formato.

[24] La tavolozza è quella di un generico modello RGB con 8 bit/canale, per un totale di $2^{8+8+8} = 2^{32} \approx 16.8$ milioni di colori.

[25] Nel GIF la trasparenza è rappresentata da uno dei 256 colori della tavolozza, che può essere indicato come tale nei metadati interni.

[26] Le GIF animate sono ancora vincolate ad avere un massimo di 256 colori per fotogramma (inclusa la trasparenza), anche se i colori scelti possono cambiare da fotogramma a fotogramma.

16. A causa della sua obsolescenza, se ne sconsiglia l'uso per nuove immagini, salvo in cui l'immagine da archiviare sia destinata all'uso in pagine web e rientri, nella sua forma originaria (o formato di partenza) all'interno dei vincoli sopraelencati.

EXR	FORMATO DI FILE
Nome completo	OpenEXR™
Estensione/i	.exr
Magic number	v/1 0x01
Tipo MIME	image/x-exr
Sviluppato da	Industrial Light and Magic
Tipologia di standard	aperto, estendibile, *de facto*, binario
Livello metadati	4
Derivato da	–
Revisione	2013
Riferimenti	• www.openexr.com/documentation • WetaDigital, *The Theory of OpenEXR Deep Samples*, 2013
Conservazione	Sì, senza compressione
Racc. per la lettura	Speciale; raccomandato in ambito cinetelevisivo
Racc. per la scrittura	Speciale; raccomandato in ambito cinetelevisivo

17. Il formato OpenEXR è stato recentemente inventato ed è oggi mantenuto, dalla comunità dei creativi dell'animazione cinematografica ma, a causa della sua versatilità e del supporto per un numero qualsiasi e facilmente estendibile di metadati, è oggi fortemente consigliato per la produzione di immagini raster –sia finali che, soprattutto, semilavorati– per la grafica e la produzione cinetelevisiva. A causa del formato specifico, non è richiesto il suo riconoscimento al di fuori delle organizzazioni dei sopracitati settori.
18. OpenEXR supporta immagini raster con le seguenti caratteristiche:

- segmentazione dell'immagine in vari modi diversi (per linee, a tasselli, per strati);
- supporto di più spazi-colore, inclusi astratti con un numero qualsivoglia elevato di canali (ciascuno associato ai propri metadati);
- supporto per un numero qualsivoglia elevato di livelli, ciascuno dotato dei propri metadati;
- data, ora ed eventuale posizione geografica della creazione;
- nome o altre caratteristiche dell'applicazione creatrice, ovvero dell'autore (se umano);
- eventuale supporto di vari algoritmi di compressione (sia *lossy* che *lossless*);
- supporto di metadati interni la cui tassonomia rientra in *namespace* estendibili.

19. Inoltre, grazie alla versatilità nel numero di livelli e di canali di colore (e la possibilità di identificarli ciascuno con un nome), è possibile memorizzare nei punti di un'immagine raster salvata in OpenEXR non soltanto il colore, quanto piuttosto informazioni relative ad altre caratteristiche tecniche, fisiche o cinetiche dell'immagine, rappresentabili punto per punto. Esempi sono:

- il cosiddetto "deep EXR", ove più livelli sovrapposti mantengono una dei medesimi oggetti a diversi gradi di "profondità" rispetto ad immagini piane;
- le *texture*, che sono comuni raster, le cui coordinate cartesiane sono chiamate (*u*,*v*), che rappresentano una "pelle" da avvolgere intorno ad una superfice artificiale tridimensionale (come quelle rappresentate in formati descritti in §2.7) mediante una trasformazione conforme che prende il nome, appunto, di mappa *uv*.

20. In alcuni casi, l'immagine tradizionalmente costituita da un raster ove i punti rappresentano dei colori è contenuta in un primo livello, mentre in altri livelli sono rappresentate altre quantità relative alla medesima immagine. Due esempi, estremamente riduttivi, di tali quantità, utili particolarmente nella produzione di effetti speciali digitali (VFX) e nell'animazione in CG (dall'inglese, "computer-grafica"), sono:

- *velocity map*, ove i 2 o 3 canali (*u*,*v*,*w*) associati ad ogni punto rappresentano le componenti cartesiane della velocità istantanea del punto della superfice;
- *optical map*, ove i canali (uno o più) associati ad ogni punto rappresentano caratteristiche ottiche del materiale in quel punto, quali ad esempio opacità/trasparenza, riflettanza, assorbanza, emittanza, ecc.

21. Per quanto riguarda i metadati interni, purtroppo, la mancanza di un registro unificato è causa del proliferare incontrollato degli stessi: unica pecca per un formato altrimenti validissimo e fortemente consigliato per la produzione di *qualsiasi* tipo di immagine raster.

JPEG2000	FORMATO DI FILE
Nome completo	JPEG 2000
Estensione/i	.png
Magic number	0x89 PNG 0x0D0A1A0A
Tipo MIME	image/jp2
Sviluppato da	Joint Photographic Experts Group
Tipologia di standard	aperto, estendibile, *de iure*, binario
Livello metadati	3
Derivato da	–
Revisione	2020/1

Riferimenti	• RFC-3745 Famiglia di standard **15444** della ISO: • ISO 15444-1:2016 • ISO 15444-2:2004 • ISO 15444-6:2013 • ISO 15444-6:2011 • ISO 15444-12:2015
Conservazione	No
Racc. per la lettura	Speciale; raccomandato nell'elaborazione professionale di immagini in vari ambiti
Racc. per la scrittura	Speciale; raccomandato nell'elaborazione professionale di immagini in vari ambiti

22. Il formato JPEG2000 utilizza una compressione basata su trasformata *wavelet*, cosiddetta multi-risoluzione, poiché l'evidenza informatica ivi rappresentata codifica prima i dettagli "grossolani" (cd. a bassa frequenza) dell'immagine, poi i dettagli via via più "fini" (cd. ad alte frequenze). A fronte di una maggiore complessità nella produzione (i.e. nella compressione) dell'immagine, il vantaggio tecnologico offerto da questo formato risiede nel fatto che la decodifica successiva dell'immagine (i.e. la sua decompressione) può avvenire anche non integralmente e interrompersi a un certo livello. In tal caso viene riprodotta una versione "a banda ristretta", cioè a risoluzione inferiore dell'immagine originale, in quanto sono state usate componenti a bassa frequenza. L'usodel formato è dunque preferibile laddove l'utilizzo professionale sia vincolato da uno o più dei seguenti fattori:

- immagini di dimensioni molto grandi, comparate a
- basse capacità computazionali del dispositivo di visualizzazione;
- canale di comunicazione tra lo storage ove è archiviata l'immagine e il dispositivo di visualizzazione.

DICOM	FORMATO DI PACCHETTO
Nome completo	Digital Imaging and Communications in Medicine
Estensione/i	
Magic number	128 byte qualsiasi, seguiti da DICM
Tipo MIME	image/dicom
Sviluppato da	National Electrical Manufacturers Association
Tipologia di standard	aperto, estendibile, *de iure*, binario
Livello metadati	2
Derivato da	–
Revisione	2018
Riferimenti	• dicom.nema.org/medical/dicom, PS3.10 • dicom.nema.org/medical/dicom, capitolo 7
Conservazione	Sì, se lo sono tutti i file del pacchetto

Racc. per la lettura	Speciale; raccomandato in ambio diagnostica sanitaria
Racc. per la scrittura	Speciale; raccomandato in ambio diagnostica sanitaria

23. Il formato DICOM consente di aggiungere alle immagini raster diversi metadati relativi all'aspetto sanitario del loro contenuto, quali ad esempio un elenco di metadati facoltativi e non esaustivi comprende:

- il nome degli specialisti che hanno eseguito l'acquisizione, l'elaborazione o la refertazione,
- il nome e modello delle apparecchiature di acquisizione e parametri elettronici, fisico-chimici, meccanici impiegati per un dato esame diagnostico,
- la data, l'ora e il nome della struttura ove sia stato effettuato l'esame diagnostico,
- eventuali commenti sul pacchetto di immagini nella sua interezza (ad esempio il testo del referto), sulla specifica immagine, o su una porzione della stessa.

DNG	FORMATO DI FILE
Nome completo	Adobe® Digital Negative
Estensione/i	.dng
Specializzazione di	TIFF/EP
Tipo MIME	image/x-adobe-dng
Sviluppato da	Adobe Systems
Tipologia di standard	proprietario (brevettato), estendibile, *de iure*, binario
Livello metadati	4
Derivato da	TIFF™
Revisione	10.1.0.0 (2017)
Riferimenti	• Adobe, *Digital Negative (DNG) Specification* v1.4.0.0, 2012
Conservazione	No
Racc. per la lettura	Speciale; raccomandato in ambito fotografico e cinetelevisivo
Racc. per la scrittura	Speciale; raccomandato in ambito fotografico e cinetelevisivo

24. Il formato DNG ("negativo digitale") di Adobe consiste in una specializzazione del formato TIFF, dotato di diverse estensioni relative alla quantità di metadati di accompagnamento (che include, tra l'altro, l'intera gamma di metadati EXIF e XMP) ma soprattutto consente di memorizzare immagini grezze ("raw") provenienti da sensori digitali con diverse caratteristiche elettroniche e geometriche. La procedura di conversione effettuata a monte permette di transcodificare un'immagine raster altrimenti memorizzata in formati di file per lo più proprietari del costruttore della fotocamera o cinepresa, mediante un applicativo che, al momento della

transcodifica, comprende il formato nativo e le informazioni sullo specifico modello dell'apparato di acquisizione. Il risultato è un'immagine visivamente fedele, identica all'originale, ma archiviata in un formato aperto e interoperabile, non più soggetto all'obsolescenza tecnologica né dell'apparato di acquisizione originario, né del formato di file proprietario di origine.

PSD	FORMATO DI FILE
Nome completo	Adobe® *Photoshop*® Standard Baseline file
Estensione/i	.psd
Magic number	8BPS
Tipo MIME	image/x-psd
Sviluppato da	Adobe Systems
Tipologia di standard	proprietario (libero), estendibile, *de facto*, binario
Livello metadati	4
Derivato da	–
Revisione	2016
Riferimenti	• www.adobe.com • FSDeveloper.com, *Adobe Photoshop File Formats Specification* 2006 • github.com/layervault/psd.rb
Conservazione	No
Racc. per la lettura	Specifico; obbligatorio in ambito beni culturali e comunicazione
Racc. per la scrittura	Specifico; obbligatorio in ambito beni culturali e comunicazione

25. Il formato nativo di Adobe *Photoshop*® si è evoluto notevolmente nel tempo e comprende attualmente la possibilità non solo di archiviare immagini a più livelli e con elevatissima variabilità tecnica, [27] ma anche informazioni vettoriali e tridimensionali di complemento alle immagini raster, che coadiuvano il loro utilizzo in molteplici contensti industriali, quali il design, la computer grafica (CG), l'architettura e la postproduzione cinetelevisiva. Si raccomanda l'utilizzo di questo formato per la creazione di semilavorati, particolarmente quando vadano elaborati dal particolare software applicativo in oggetto. Per l'archiviazione a lungo termine e la conservazione invece, si consiglia di riversare il contenuto in altri formati più aperti e interoperabili (quali OpenEXR, DNG o TIFF).

ARRIRAW	FORMATO DI FILE
Nome completo	ARRIRAW
Estensione/i	.ari
Magic number	ARRI 0x12345678; ARRI 0x78563412

[27] Si segnala la presenza un dialetto del formato PSD specificatamente per archiviazione e riutilizzo di immagini di grandissime dimensioni: il *Photoshop Large Format* (formato molto simile all'originario, che adotta l'estensione di file .psb).

Tipo MIME	image/arriraw
Sviluppato da	Arnold & Richter Cine Technik GmbH
Tipologia di standard	proprietario (libero), estendibile, *de iure*, binario
Livello metadati	4
Derivato da	–
Revisione	2018
Riferimenti	• documentazione registrata SMPTE RDD30:2014 • documentazione registrata SMPTE RDD31:2014
Conservazione	No
Racc. per la lettura	Speciale; riprese cinematografiche
Racc. per la scrittura	Speciale; sconsigliata la produzione al di fuori dell'ambito cinematografico

26. Il formato ARRIRAW è un formato proprietario della ARRI GmbH, ma aperto, utilizzato attualmente come formato nativo "*raw*" per le sue cineprese digitali, che è stato candidato per diventare standard SMPTE. Nonostante si tratti di un formato proprietario, ne è un esempio virtuoso, in quanto esso supporta più tipologie di codifiche di immagine, nonché di metadati avanzati a contorno dell'immagine stessa, quali le informazioni colorimetriche (codifica dello spazio-colore, parametri di correzioni impostati direttamente sulla cinepresa, importati *una tantum* da file, ovvero comunicati in tempo reale da un'applicazione con connessione di rete diretta sulla cinepresa), quelle relative all'area attiva del fotogramma, alla velocità di ripresa mediante TimeCode (il formato ne supporta fino a 7 tipi differenti e indipendenti fra loro), ai metadati di produzione, ecc. In teoria, il formato ARRIRAW potrebbe essere utilizzato per produrre nuove immagini raster in una miriade di codifiche e di formati.

DPX	FORMATO DI FILE
Nome completo	Digital Picture Exchange
Estensione/i	.dpx
Magic number	SDPX, XPDS
Tipo MIME	image/x-dpx
Sviluppato da	Society of Motion Picture and Television Engineers
Tipologia di standard	aperto, estendibile, *de iure*, binario
Livello metadati	4
Derivato da	Kodak® Cineon™
Revisione	2.0 (2014)
Riferimenti	• standard SMPTE ST268:2014
Conservazione	No
Racc. per la lettura	Specifico; obbligatorio in ambito cinetelevisivo

Racc. per la scrittura	Specifico; nessuna raccomandazione

27. Il formato DPX (originariamente sviluppato da Kodak come evoluzione del formato Cineon® usato nell'omonimo scanner digitale per le pellicole cinematografiche 35mm e successivamente divenuto uno standard SMPTE) permette la rappresentazione di immagini raster con differenti caratteristiche tecniche, oltre a un considerevole numero di metadati. L'attuale versione 2.0 suddivide i metadati in quattro categorie: generici, cinematografici, televisivi e "utente". Tra i metadati delle categorie cinetelevisive vi è la possibilità di rappresentare indicazioni per collocare con precisione l'immagine come fotogramma di una particolare sequenza video, quali:

- numerazione del fotogramma relativa all'inizio della sequenza,
- TimeCode (con relativi *framerate* e modalità di scansione del fotogramma, cfr. §2.10),
- KeyKode™ del forogramma rispetto al rullo di pellicola fotochimica da cui è stato scansionato digitalmente (con relativa *film-speed*).

28. Il formato si presta dunque ad una rappresentazione di documenti video mediante pacchetti di file, ove ciascun fotogramma è archiviato in un file separato (cfr. pacchetto DI basato su DPX, cfr. §2.12). In linea di principio il formato DPX ammette tra i metadati anche quelli che associano colorimetria digitale rappresentata nel raster dalla densitometria dell'emulsione di partenza (da cui il fotogramma è stato scansionato) o di destinazione (sulla quale s'intende possa essere stampato digitalmente), anche se pochissimi applicativi per la postproduzione cinematografica ne fanno realmente uso. Nonostante il DPX sia stato l'indiscusso formato di riferimento per la produzione, postproduzione e archiviazione digitale di contenuti cinematografici sin dall'avvento del Digital Intermediate (*DI*, cfr. §2.12), tre fattori ne hanno rapidamente ridotto l'importanza:

- il declino della produzione cinematografica su pellicola e la consistente riduzione di processi di scansione digitale di tali pellicole (ora quasi esclusivamente effettuati a scopo di restauro e conservazione);
- il mancato utilizzo di moltissimi metadati rappresentabili nel formato DPX da parte di gran parte dei software applicativi, che avrebbe invece incrementato l'efficacia di tale formato;
- la comparsa di nuovi standard tecnologici per la rappresentazione dell'immagine digitale, quali ad esempio l'elevata latitudine di posa (*HDR*). A tale scopo l'SMPTE sta sviluppando una nuova versione più moderna del formato DPX.

29. Mentre si obbliga agli enti tecnici operanti nel settore della postproduzione, archiviazione e conservazione cinematografica di leggere immagini nel formato DPX, allo scopo di preservare l'accessibilità a contenuti provenienti da archivi in pellicola già digitalizzati o archivi digitali di DI basati su DPX, si raccomanda fortemente di considerare formati alternativi, quali ad esempio OpenEXR, ovvero i pacchetti di master interoperabile (IMF, cfr. §2.12).

ACES	
Nome completo	Academy Color Encoding System
Estensione/i	.exr, .mxf, .amf, .clf
Tipologie MIME	image/exr, application/mxf
Sviluppato da	Academy of Motion Picture Arts and Sciences
Tipologia di standard	aperto, estendibile, retrocompatibile, *de iure*
Derivato da	–
Revisione	1.1.0 (2018)
Riferimenti	Standard e bollettini tecnici dell'AMPAS: • TB-2014-012, *ACES version 1.0 component names* Famiglia di standard **ST2065** della SMPTE: • ST2065-1:2012, *Academy Color Encoding Specification* • ST2065-4:2013, *ACES image container file layout* • ST2065-5:2016, *mapping ACES into MXF container* • ST2067-50:2018, *IMF – Application #5: ACES* • www.oscars.org/aces • www.acescentral.com • W. Arrighetti, 'The ACES: a professional color-management framework for production, post-production and archival of still and motion pictures', *Journal of Imaging*, Vol.3, №4, pp.189–225, MDPI, 2017
Conservazione	Sì
Racc. per la lettura	Speciale; consigliato in post-produzione cinetelevisiva
Racc. per la scrittura	Speciale; consigliato in post-produzione, archiviazione e conservazione di contenuti cinetelevisivi in ACES

30. Il sistema *Academy Color Encoding System* (*ACES*), standard SMPTE introdotto dall'AMPAS, permette di descrivere in maniera unificata gli spazi-colore e i processi produttivi di elaborazione e trasporto del colore digitale nelle applicazioni professionali in campo cinetelevisivo. ACES si prefigge come una possibile soluzione alla rappresentazione di immagini raster (e video) catturate da dispositivi

e camere con molteplici colorimetrie (spesso proprietarie), al trasporto di metadati colorimetrici tra processi dominati da svariati *vendor*, e visualizzabili con una moltitudine di tecnologie di visualizzazione (p.es. LED, OLED, plasma, ecc., così come HDR+, Dolby Vision,® HLG, ecc.). Il trasporto di file codificati con la colorimetria ACES può avvenire sia con il formato OpenEXR con l'aggiunta di specifici vincoli e metadati[28] (standard ST2065-4), sia imbustando una sequenza di fotogrammi, già nel formato EXR, all'interno di un contenitore MXF (cfr. standard ST2065-5 e §2.12) anch'esso con determinati vincoli e metadati. Ulteriori informazioni colorimetriche in ACES possono essere rappresentate accompagnando, il contenuto audiovisivo in un pacchetto di file contenente uno o più *sidecar file* nel formato AMF (cfr. §2.12).

2.6.1 Raccomandazioni per la produzione di documenti

1. Come già discusso sopra all'inizio della sezione, i formati di immagini raster vanno distinti innanzi tutto tra le finalità generiche e specializzate — queste ultime differenti a seconda dell'ambito: sanità, architettura e urbanistica, grafica e pubblicità (inclusa la produzione online), intrattenimento (inclusi l'animazione o CG, così come produzione, postproduzione ed effetti speciali cinetelevisivi).
2. Per i formati generici (PNG, JPEG, TIFF, GIF) sussiste in generale l'obbligo di riconoscimento. Nonostante tale obbligo in merito ai formati generali, le singole PP.AA., qualora accettino documenti informatici in formati di immagini raster generici tra quelli sopra elencati per un loro procedimento amministrativo o a scopo di conservazione, possono decidere di limitare ulteriormente l'accettazione di tali formati ad un loro sottoinsieme, specificatamente allo scopo di ridurre o uniformare i formati dei documenti.
3. 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,

[28] Il vincolo principale è nell'assenza di compressione e nella codifica nello spazio-colore ACES2065-1 (standard ST2065-1); devono inoltre essere presenti alcuni metadati colorimetrici facoltativi del formato OpenEXR, più un metadata booleano specifico di ACES.

livelli aggiuntivi e un certo livello di "metadati tecnici" (e.g. spazio-colore, 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.

4. Per quanto riguarda settori specifici, si raccomanda invece l'utilizzo di formati specializzati nei settori produttivi di riferimento, quali editoria, grafica, pubblicità, sanità, edilizia, architettura, intrattenimento, produzione cinetelevisiva. Si raccomandano perciò le seguenti adozioni:

- Il formato OpenEXR (anche detto, colloquialmente, "EXR") per immagini raster relative a produzione, postproduzione, archiviazione e soprattutto conservazione di contenuti cinetelevisivi, pubblicitari, videoludici, animazione e beni culturali — che si tratti di semilavorati, master o

documenti definitivi. Il formato sia inoltre corredato adeguatamente dal più ampio spettro possibile di metadati a scopo archivistico.

- Formato DICOM per tutte le immagini diagnostiche di provenienza sanitaria, adeguatamente corredato dai metadati di produzione e refertazione,
- 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.
- Formato DPX per fotogrammi provenienti direttamente da scansioni digitali di pellicole cinematografiche non ulteriormente elaborate (per le quali si preferisce un riversamento nel formato OpenEXR durante le fasi della postproduzione).
- Formato ARRIRAW per le immagini provenienti dal girato originale di cinecamere ARRI (quali quelle della famiglia ALEXA ovvero AMIRA) — salvo quando esigenze di produzione cinetelevisiva richiedano che le camere *formino* il loro girato direttamente in formati video che adottano compressione *lossy*. [29]
- Formato PSD per immagini definitive provenienti da, ovvero semilavorati destinati all'elaborazione da parte di software di fotoritocco specializzati che supportino tale formato — nativamente o per compatibilità.

5. Tra i formati sopra individuati, il PNG e le varianti di TIFF e OpenEXR (EXR) 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).

[29] Ad esempio, contenitori QuickTime o MXF (§2.12) che imbustano video Apple® ProRes® ovvero Avid® DNxHD™, cfr. §2.10.

2.7 Immagini vettoriali e modellazione digitale

1. Si chiamano vettoriali le immagini rappresentate nel file mediante una descrizione algoritmica della geometria e del colore che le compone, utilizzando tecniche descrittive più o meno complesse. I modelli multidimensionali sono una generalizzazione delle immagini vettoriali utilizzando tre o più dimensioni (a prescindere dal problema della visualizzazione, che è spesso demandato alle applicazioni più o meno specializzate).
2. La descrizione delle immagini vettoriali è fatta in uno spazio geometrico virtuale, mediante coordinate tipicamente adimensionali, sebbene nei metadati interni di alcuni tipi di file impiegati, e/o nelle informazioni metriche usate dalle primitive, possono essere presenti riferimenti a dimensioni reali. Anche la rappresentazione di caratteristiche quali colore, trasparenze o riflettanze può avvenire mediante una o più delle seguenti tecniche:

- coordinate di spazi-colore riferiti ad ambienti a luce emessa o riflessa (in inglese, rispettivamente, *output-referred* o *scene-referred*), ovvero ad altri tipi di colorimetrie.
- etichette di colore che rimandano a specifici cataloghi di vernici o altri coloranti (p.es. "blu cobalto", "Pantone®33M", "smalto dorato", ecc.),
- etichette che rimandano a cataloghi di materiali, tipicamente specifici per applicazione o riferiti altrove mediante file o etichette che rimandano a basi di dati esterne (p.es. "mogano laminato", "calcestruzzo", "cristallo smerigliato", "ebano lucido", "lino", "tartan", ecc.).

3. La potenziale multidimensionalità rappresentata in questi file, e la loro caratteristica intrinseca di descrivere immagini scomposti in oggetti distinti (tipicamente primitive geometriche con caratteristiche aggiuntive quali nomi/etichette, colori, trasparenze, collegamenti logici ad altri oggetti, ecc.) consente di salvare anche informazioni circa l'animazione –automatica o programmabile– di alcuni oggetti rispetto ad altri.
4. I formati di file più usati e specifici per immagini vettoriali sono: SVG, Adobe® *Illustrator*® (.ai), Encapsulated PostScript™ (.eps).
5. I modelli digitali sono un'estensione delle immagini vettoriali (cfr. §2.7) ma a spazi con tre o più dimensioni. Anche in questo caso gli oggetti virtuali sono descritti in maniera geometrica e corredati da metadati. Le applicazioni principali per questi file sono la progettazione edile e industriale, l'architettura, l'intrattenimento e le scienze in genere.
6. I formati di file più usati per modelli bi- e tridimensionali sono ben distinti tra categorie aperte e categorie chiuse (sia rispetto alle specifiche che rispetto alla proprietà intellettuale — cfr. §2.7):

Tra i primi ci sono Autodesk® DXF, Autodesk® FBX, Blender (.blend), Wavefront OBJ, Google SketchUp (.skp), Stereolitography (.stl); tra i secondi, invece, figurano Autodesk® *AutoCAD*® (.dwg), Autodesk® *Maya*® (.ma, .mb), Autodesk® *3ds Max*® (.3ds), Maxon® *Cinema4D* (.c4d).

SVG	FORMATO DI FILE
Nome completo	Scalable Vector Graphics
Estensione/i	.svg, .svgz
Specializzazione di	XML (*namespace* svg)
Tipo MIME	image/svg+xml, image/svg+xml+zip
Sviluppato da	World Wide Web Consortium
Tipologia di standard	aperto, estendibile, *de jure*, testuale
Livello metadati	4
Derivato da	–
Revisione	1.1 (second edition)
Riferimenti	• W3C Recommendation *SVG 1.1 (2nd Ed.)*, 16 agosto 2011 • github.com/W3C/svgwg
Conservazione	Sì; cfr. §2.8
Racc. per la lettura	Generico; obbligatorio
Racc. per la scrittura	Generico; fortemente raccomandato

7. Il formato SVG, basato su XML, descrive un'immagine vettoriale componente per componente, utilizzando l'estendibilità del linguaggio per "etichettare" opzionalmente alcune parti di queste componenti affinché possano essere referenziate da altri documenti o applicativi (ad es., una pagina web in HTML e il suo foglio di stile CSS allo scopo di migliorare l'interattività con le immagini vettoriali). Data la sua versatilità e l'apertura dello standard è il formato fortemente raccomandato per l'uso, la trasmissione e la conservazione di tutte le immagini vettoriali.
C'è una nuova versione proposta per essere la futura SVG 2.0; essa supporta alcune caratteristiche quali l'animazione dei componenti. Tale versione, tuttavia, non è allo stato attuale raccomandata.

ILLUSTRATOR	FORMATO DI FILE
Nome completo	Adobe® *Illustrator*® artwork
Estensione/i	.ai
Specializzazione di	PDF, Encapsulated PostScript™
Tipo MIME	application/illustrator
Sviluppato da	Adobe Systems
Tipologia di standard	proprietario (libero), *de facto*, binario
Livello metadati	3

Derivato da	Adobe® Encapsulated PostScript™, Adobe® PDF
Revisione	2019
Riferimenti	• Adobe Developer Support, *Adobe Illustrator File Format Specification*, 1998
Conservazione	No; cfr. §2.8
Racc. per la lettura	Specifico; nessuna raccomandazione
Racc. per la scrittura	Specifico; non raccomandato salvo che per scambio di semilavorati (a breve termine) tra reparti di design

8. Il formato proprietario dell'applicativo di disegno vettoriale Adobe® *Illustrator*® è in realtà un contenitore di un tipo specializzato di documento PDF ovvero di Encapsulated PostScript™, contenente estensioni particolari interpretabili dall'applicativo. Il formato è evoluto parecchio con le versioni successive di *Illustrator*®, che però ha sempre mantenuto la retrocompatibilità totale. Essendo l'applicativo molto diffuso tale formato è qui indicato perché è uno standard di riferimento per documenti in particolari ambiti (editoria e grafica pubblicitaria). Mentre è un formato consigliato per dei semilavorati (qualora vadano elaborati tramite *Illustrator*®), si sconsiglia di utilizzarlo per la produzione di documenti destinati alla conservazione, valutando piuttosto un riversamento dei documenti esistenti in altro formato più interoperabile (e.g. SVG).

ENCAPSULATED POSTSCRIPT	FORMATO DI FILE
Nome completo	Encapsulated PostScript®
Estensione/i	.eps
Specializzazione di	PostScript®
Tipo MIME	image/eps, application/eps
Sviluppato da	Adobe Systems
Tipologia di standard	proprietario (libero), *de facto*, binario, deprecato
Livello metadati	1
Derivato da	PostScript®
Revisione	3.0 (1992)
Riferimenti	• Adobe, *Encapslutated PostScript File Format Specifications*, v3.0 (1992)
Conservazione	No; cfr. §2.8
Racc. per la lettura	Specifico; raccomandato nei reparti di design
Racc. per la scrittura	Specifico; non raccomandato

9. Il formato EPS (Encapsulated PostScript™) è un dialetto del linguaggio PostScript™ (cfr. §2.1) con alcune distinzioni, come ad esempio la rappresentazione di un oggetto all'intero di una regione rettangolare confinata in una singola pagina

e la possibilità di memorizzare una "pre-visualizzazione" raster (cd. *preview*) dell'intero contenuto. Di fatto, il formato è utilizzato il più delle volte per rappresentare singoli disegni o immagini da allegare successivamente a documenti PostScript™ impaginati ovvero, per via dell'apertura del formato, importarlo in altri applicativi di elaborazione vettoriale. Dato che il contenuto del documento è descritto nel linguaggio omonimo, orientato alle primitive di stampa, i file EPS non descrivono le immagini vettoriali come "oggetti" nell'accezione moderna, non supportando scenari di utilizzo più modulari o estendibili, come invece permette la scelta di formati più evoluti, quali ad esempio l'SVG. Il formato Encapsulated PostScript™ è dunque raccomandato in lettura per aprire documenti già esistenti, ma si sconsiglia la produzione di altri documenti in tale formato, invitando inoltre le PP.AA. a valutare il riversamento dei file esistenti in altri formati interoperabili (e.g. SVG).

ODG	FORMATO DI FILE
Nome completo	Open Document Graphics
Estensione/i	.odg
Magic number	XML imbustato dentro ZIP
Tipo MIME	application/vnd.oasis.opendocument.graphics
Sviluppato da	Organization for the Advancement of Structured Information Standards
Tipologia di standard	aperto, estendibile, *de iure*, binario
Livello metadati	3
Derivato da	—
Revisione	1.2 (2015)
Riferimenti	Famiglia di standard **26300** della ISO/IEC: • ISO/IEC 26300-1:2015, *ODF for Office Applications v1.2 – Part 1: OpenDocument Schema* • ISO/IEC 26300-3:2015, *ODF for Office Applications v1.2 – Part 3: Packages* • OASIS, *Open Document Format for Office Applications (OpenDocument)*, v1.2 (2015)
Conservazione	No; cfr. §2.8
Racc. per la lettura	Generico; raccomandato solo a scopo di riversamento
Racc. per la scrittura	Sconsigliato

10. Il formato ODG è il dialetto del formato OpenDocument (cfr. §2.3) per le immagini vettoriali; è incluso in questo elenco prevalentemente perché è completamente aperto ed è un necessario complemento agli altri dialetti, anche se il suo principale impiego è come allegato presente all'interno di testi impaginati (.odt), fogli di calcolo (.ods) e presentazioni multimediali (.odp).

DXF	FORMATO DI FILE
Nome completo	AutoCAD® Drawing Interchange Format
Estensione/i	.dxf
Magic number	– / AutoCAD Binary DXF 0x0D0A1A00
Tipo MIME	image/vnd.dxf
Sviluppato da	Autodesk
Tipologia di standard	proprietario, libero, *de facto*, testuale e binario
Livello metadati	2
Derivato da	–
Revisione	2020/1
Riferimenti	• Autodesk, *DXF Reference*, 2011
Conservazione	No; cfr. §2.8
Racc. per la lettura	Specifico; nessuna raccomandazione
Racc. per la scrittura	Specifico; sconsigliato in archiviazione/conservazione

11. Pur trattandosi di un formato testuale (che descrive gli oggetti che compongono il modello o l'immagine mediante *primitive grafiche*) DXF permette una compressione facoltativa allo scopo di ridurre l'occupazione digitale del file. Tuttavia il numero di primitive grafiche utilizzate in questo formato non permette una rappresentazione con il medesimo livello semantico o di dettaglio tecnico di altri formati (p.es. i colori o i materiali); di conseguenza il formato non è completamente interoperabile con essi o con i loro applicativi.

DWF	FORMATO DI FILE
Nome completo	AutoCAD® Design Web Format
Estensione/i	.dwfx, .dwf
Magic number	–
Tipo MIME	model/vnd.dwf, drawing/dwf, image/dwf
Sviluppato da	Autodesk
Tipologia di standard	aperto, estendibile, *de iure*, testuale
Livello metadati	3
Derivato da	Open Packaging Convention
Revisione	6.0
Riferimenti	• Autodesk Knowledge Network, *About DWF and DWFx Files* • ISO/IEC 29500-2:2012
Conservazione	No; cfr. §2.8
Racc. per la lettura	Specifico; raccomandato in applicazioni CAD

Racc. per la scrittura	Specifico; raccomandato per scambio e archiviazione di disegni e modelli tecnici in ambito CAD

12. Il formato DWF è stato specificatamente pensato per l'interoperabilità; è infatti basato, dalla versione 6.0, sullo standard ISO/IEC 29500-2 dell'*open packaging* (lo stesso usato da OOXML, §2.3), e dunque costituito da un file ZIP contenente più file differenti (XML per la descrizione del modello; PNG per le *preview* e le *texture*). Pur trattandosi di un formato testuale (che descrive gli oggetti che compongono il modello o l'immagine mediante *primitive grafiche*) il medesimo formato permette anche la compressione allo scopo di ridurre l'occupazione digitale del file. È spesso utilizzato come formato di interscambio e, tra tutti i formati imparentati, è quello con le specifiche più aperte, perciò è utilizzabile anche per archiviazione e conservazione.

DWG	FORMATO DI FILE
Nome completo	Autodesk® AutoCAD® Drawing
Estensione/i	.dwg, .dwt
Magic number	AC1032 (le ultime cifre variano con versione)
Tipo MIME	application/acad, image/vnd.dwg
Sviluppato da	Autodesk
Tipologia di standard	proprietario, libero, *de facto*, binario
Livello metadati	2
Derivato da	—
Revisione	DWG 2018
Riferimenti	• Autodesk, *AutoCAD .dwg file format* • Open Design, *Specification for .dwg*, v5.4.1 (2018)
Conservazione	Sì; cfr. §2.8
Racc. per la lettura	Specifico; nessuna raccomandazione
Racc. per la scrittura	Specifico; sconsigliato in archiviazione/conservazione

13. Il formato DWG è anch'esso proprietario ma ampiamente utilizzato per lo scambio di immagini vettoriali (prevalentemente bidimensionali) nel campo del disegno tecnico e dell'architettura.

FBX	FORMATO DI FILE
Nome completo	Autodesk® FBX®
Estensione/i	.fbx
Magic number	AC1032
Tipo MIME	model/vnd.fbx
Sviluppato da	Autodesk

Tipologia di standard	proprietario (libero), chiuso, estendibile, *de facto*
Livello metadati	3
Derivato da	–
Revisione	2019 Extension 2
Riferimenti	• Autodesk Knowledge Network, *Autodesk FBX Files* • Blender Foundation, *FBX Binary file format specification*
Conservazione	Sì; cfr. §2.8
Racc. per la lettura	Specifico; raccomandato per modelli di progettazione
Racc. per la scrittura	Specifico; sconsigliato in archiviazione/conservazione

14. Il formato FBX, proprietario e chiuso ma completamente libero nel suo utilizzo, è utilizzato da vari applicativi di modellazione tridimensionale (ambito CAD e altro) prevalentemente come formato di interscambio. La sua estrema versatilità e il supporto del formato da parte di diversi applicativi professionali lo renderebbero adatto anche per archiviazione e conservazione, se non fosse per l'assenza di specifiche tecniche condivise.

STL	FORMATO DI FILE
Nome completo	Stereolithography file format
Estensione/i	.stl
Magic number	solid; –
Tipo MIME	model/stl, model/x.stl-ascii|binary
Sviluppato da	3D Systems (Albert-Battaglin Consulting Group)
Tipologia di standard	proprietario (libero), estendibile, *de facto*
Livello metadati	1
Derivato da	–
Revisione	2.0 (2009)
Riferimenti	• J.D. Hiller, H. Lipson, *STL 2.0: a proposal for a universal multi-material additive manufacturing file format*, 2009 • all3dp.com/what-is-stl-file-format-extension-3d-printing
Conservazione	No
Racc. per la lettura	Specifico; obbligatorio nel campo della stampa 3D
Racc. per la scrittura	Specifico; raccomandato nel campo della stampa 3D

15. Il formato STL (da non confondersi con l'omonimo formato per sottotitoli, cfr. § 2.11) è uno dei tanti standard *de facto* ed interoperabili per la modellazione

tridimensionale industriale; recentemente si è affermato come formato d'elezione per la trasmissione di modelli per produzione tramite le tecniche di "stampa 3D" o stereolitografia. In quanto formato completamente aperto si consiglia il suo utilizzo per la produzione di prototipi con tale destinazione d'uso o loro archiviazione, anche se la mancanza di caratteristiche più avanzate (adatte alla prototipazione professionale (e.g. in ambito ingegneristico) non lo rendono un candidato ideale per qualunque modello tridimensionale.

2.7.1 Raccomandazioni per la produzione di documenti

1. 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.

2.8 Caratteri tipografici

1. Un esempio specializzato di immagine vettoriale –in questo caso bidimensionale e monocromatica– è costituita dal singolo "glifo" (i.e. carattere tipografico), per il quale possono essere rappresentate forme distinte nello stesso file. All'interno di un file di fonti tipografiche (*font* in inglese) esiste una tabella ove, per un sottoinsieme di tutti i possibili caratteri tipografici rappresentabili da un sistema operativo (e codificati mediante codici numerici negli standard ASCII, UNICODE, ovvero UTF-8 e UTF-16), è contenuta la rappresentazione vettoriale del glifo corrispondente al carattere. In un file di font ai caratteri può corrispondere più di un glifo, che rappresenta varianti di quel carattere a seconda del tipo (e.g. tondo o corsivo), del peso (e.g. sottile, regolare, grassetto, nero), di altri fattori stilistici (e.g. numerali vecchio stile o meno) e, in alcuni casi, della dimensione/grandezza del carattere per tener conto del cosiddetto "aggiustamento ottico". In alcuni casi tali variazioni stilistiche dei glifi sono mantenute in file separati (ad esempio un file per ciascuna combinazione di tipi e pesi), che vanno a costituire dunque un "pacchetto di font" (cfr. definizioni date in §1.1.2).
2. Le seguenti famiglie di caratteri tipografici –organizzati in macro-tipologie– sono considerati "standard" da diversi organi di settore (come il W3C) — si leggano le raccomandazioni più sotto:

- "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.

3. I caratteri tipografici di cui al punto precedente sono chiamati, all'interno delle Linee guida di cui questo Allegato fa parte, "font interoperabili". Le PP.AA. che producono documenti informatici il cui contenuto dipende anche dai caratteri tipografici impiegati, si impegnano a prendere le misure tecnico-organizzative del caso (inclusa l'effettuazione di una valutazione di interoperabilità, cfr. §3.1), affinché la produzione di tali documenti includa i caratteri tipografici utilizzati nel file stesso ovvero sia costituito da un pacchetto di file contenente anche i caratteri tipografico. Qualora ciò non sia tecnicamente possibile, le PP.AA. si adoperano per utilizzare solamente caratteri tipografici interoperabili.
4. I formati più diffusi e interoperabili di font tipografici sono

d) TrueType (.ttf) e la sua evoluzione, OpenType (.otf), utilizzati negli applicativi dei principali SO tradizionali e mobile;
e) Web Open Font Format (.woff, .woff2), usati per le pagine web.

OpenType		Formato di file
Nome completo	OpenType®	
Estensione/i	.otf	
Magic number	OTTO 0x00	
Tipo MIME	font/otf, application/x-font-otf	
Sviluppato da	Microsoft, Adobe, Apple, Google	
Tipologia di standard	proprietario (libero), *de iure*, binario	
Livello metadati	3	
Derivato da	TrueType®	
Revisione	1.8.3	
Riferimenti	• ISO/IEC 14496-22:2015 • Microsoft OpenType® specifications 1.8.3 (2018) • RFC-2361	
Conservazione	Sì	
Racc. per la lettura	Generico; obbligatorio	
Racc. per la scrittura	Specifico; obbligatorio	

TrueType		Formato di file
Nome completo	TrueType®	
Estensione/i	.ttf	

Magic number	true 0x00
Tipo MIME	font/ttf, application/x-font-ttf
Sviluppato da	Apple, Microsoft
Tipologia di standard	proprietario (libero), *de facto*, binario
Livello metadati	1
Derivato da	Adobe Type-1
Revisione	1994
Riferimenti	• Apple Developer, *TrueType™ Reference Manual* • Microsoft Typography webpage
Conservazione	Sì
Racc. per la lettura	Generico; obbligatorio
Racc. per la scrittura	Specifico; valutare utilizzo del formato OpenType

WOFF	FORMATO CONTENITORE
Nome completo	Web Open Font Format
Estensione/i	.woff2; .woff
Magic number	wOF2; wOFF
Tipo MIME	font/woff2, application/font-woff
Sviluppato da	Mozilla Foundation, Opera Software, Microsoft
Tipologia di standard	proprietario (libero), *de iure*, binario
Livello metadati	1
Derivato da	OpenType®
Revisione	2.0
Riferimenti	• W3C Recommendation, *WOFF 2.0*, 2018 • W3C Recommendation, *WOFF 1.0*, 2012 • github.com/W3C/woff
Conservazione	Sì
Racc. per la lettura	Specifico; raccomandato per tipografia di contenuti web
Racc. per la scrittura	Specifico; raccomandato per caratteri nonstandard in pagine web

2.8.1 Raccomandazioni per la produzione di documenti

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

a) Qualora di adoperino caratteri tipografici standard (come indicati all'inizo 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),[30] ovvero si formi un pacchetto di file (cfr. §1.1.1) 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 caratteti alfanumerici, di interpunzione e diacritici del linguaggio utilizzato, della lingua italiana e di quella inglese inglese.
d) Ottemperare, se possibile, alla raccomandazione 1. per i caratteri alfanumerici, di interpunzione e diacritici di tutte le lingue dell'Unione Europea;[31]
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[32] con altri caratteri standard, senza che ciò risulti in una variazione sostanziale del contenuto informativo del documento.

2. Per ogni altra considerazione riguardo l'uso dei caratteri tipografici nei documenti informatici si rimanda alle *Linee guida sull'accessibilità* e alle *Linee guida di design*, anch'esse emanate dall'Agenzia per l'Italia Digitale.

2.9 Audio e musica

1. La rappresentazione digitale dei segnali audio è divisa in due categorie: quella per "forme d'onda" e quella, più indiretta, per metadati. Nel primo caso il segnale sonoro è rappresentato mediante un'approssimazione digitale dell'onda sonora, registrata o riprodotta. In fase di registrazione del sonoro la forma d'onda si ottiene mediante campionamento (misurato in numero di campioni per secondo, o *sample rate*) e quantizzazione (misurato in bit per campione); il prodotto di tali misure

[30] Come accade, ad esempio per alcuni tipi di file PDF (§2.1) di pacchetti IMF ovvero DCP (§2.12).
[31] Le raccomandazioni 1. e 2. servono a confermare, a priori, che il carattere tipografico può potenzialmente rappresentare altri caratteri tipografici e alfabetici qualora il testo andasse modificato ovvero tradotto in altra lingua.
[32] Tali sostituzioni possono avvenire anche a catena, purché terminino sempre con un carattere tipografico standard (denominato, in inglese, *fallback* font). Si consideri, come esempio, la seguente catena di sostituzioni tipografiche: Minion Pro > Minion Std > Noto Serif > Albertina > Times New Roman > Times > serif.

costituisce il data-rate audio (anche detto *bit-rate* perché misurato in bit per secondo). Un file o un flusso audio può rappresentare le forma d'onda di più sorgenti sonore contemporaneamente — dette "canali".
2. I campioni audio di una forma d'onda possono subire una compressione digitale allo scopo di ridurne le dimensioni occupate. I formati più noti per la memorizzazione di tali segnali audio sono il Waveform RIFF (.wav), MP3, FLAC, OGG e Broadcast Wave (.bwf).
3. Nel secondo caso, invece, il file contiene una rappresentazione temporale di suoni pre-campionati (che possono essere contenuti nel medesimo file, o riferendosi a file esterni contenenti le loro forme d'onda) e degli eventuali effetti associati a tali campioni ovvero alla loro riproduzione. L'applicazione principale per tali file è la rappresentazione di brani musicali, ove il timbro di ciascuno strumento è associato a banche dati generiche di suoni (che obbediscono a standard quali General MIDI) ovvero salvato come forma d'onda esterna, mentre il file audio contiene la partitura completa, nota per nota. Appartengono a questa tipologia i file musicali registranti uno standard quale il MIDI (.mid).

WAV		FORMATO CONTENITORE
Nome completo	[Broadcast] Waveform File	
Estensione/i	.wav, .bwf, .rf64	
Magic number	WAVE, RIFF	
Tipo MIME	audio/wave	
Sviluppato da	IBM; Microsoft	
Tipologia di standard	proprietario, libero, *de iure*, binario	
Livello metadati	3	
Derivato da	Resource Interchange File Format (RIFF)	
Revisione	2018	
Riferimenti	• Microsoft, *Multimedia Data Standards Update* (1994) • RFC-2361 • EBU Recommenation R111 (2007) • EBU – Tech 3285-1, *Broadcast Wave (BWF)* (2011) • EBU – Tech 3306 , *RF64* (2018) • soundfile.sapp.org/doc/WaveFormat	
Conservazione	Sì, senza compressione	
Racc. per la lettura	Generico; obbligatorio per audio qualunque	
Racc. per la scrittura	Generico; fortemente raccomandato per audio qualunque	

MP3	FORMATO DI FILE / CODEC
Nome completo	MPEG-1, Layer 3
Estensione/i	.mp3
Magic number/FourCC	0xFFFB30; ID3 / mp3
Tipo MIME	audio/mpeg
Sviluppato da	Moving Pictures Expert Group
Tipologia di standard	aperto, estendibile, *de iure*, binario
Livello metadati	4
Derivato da	MPEG-ES
Revisione	1998
Riferimenti	• ISO/IEC 11172-3:1993 • ISO/IEC 13818-3:1998 • mpgedit.org/mpgedit/mpeg_format/MP3Format.html
Conservazione	No
Racc. per la lettura	Generico; obbligatorio
Racc. per la scrittura	Generico; raccomandato per contenuti musicali

AIFF	FORMATO DI FILE
Nome completo	Audio Interchange File Format
Estensione/i	.aiff, .aifc, .aif
Magic number	AIFF; AIFC
Tipo MIME	audio/aiff
Sviluppato da	Moving Pictures Expert Group
Tipologia di standard	proprietario (libero), estendibile, *de facto*, binario
Livello metadati	3
Derivato da	Electronic Arts® Interchange File Format (IFF)
Revisione	1.3 (1991)
Riferimenti	• Apple Developer, *Audio Interchange File Format: "AIFF"* • www-mmsp.ece.mcgill.ca/Documents/AudioFormats/AIFF
Conservazione	No
Racc. per la lettura	Nessuna raccomandazione
Racc. per la scrittura	Nessuna raccomandazione

FLAC	FORMATO DI FILE / CODEC
Nome completo	Free Lossless Audio Codec
Estensione/i	.flac

Magic number/FourCC	fLaC / FLAC
Tipo MIME	audio/flac
Sviluppato da	comunità open source
Tipologia di standard	aperto (licenze GNU GPL e BSD), *de facto*, binario
Livello metadati	3
Derivato da	–
Revisione	1.3.2 (2017)
Riferimenti	• xiph.org/flac
Conservazione	Sì
Racc. per la lettura	Specifico; raccomandato
Racc. per la scrittura	Specifico; raccomandato

RAW — FORMATO DI FILE

Nome completo	Audio "Raw"
Estensione/i	.pcm, .raw, .sam, …
Magic number	–
Tipo MIME	audio/basic
Sviluppato da	comunità open source e altri *vendor*
Tipologia di standard	aperto, *de facto*, muto, binario
Livello metadati	1
Derivato da	–
Revisione	–
Riferimenti	• www.fmtz.com/misc/raw-audio-file-formats • en.wikipedia.org/wiki/Au_file_format • www-mmsp.ece.mcgill.ca/Documents/AudioFormats/AU
Conservazione	Sì
Racc. per la lettura	Specifico; raccomandato per le registrazioni ambientali
Racc. per la scrittura	Specifico; raccomandato per le registrazioni ambientali

VORBIS — CODEC

Nome completo	Vorbis
Profili	–
Codice FourCC	vorb
Sviluppato da	comunità open source
Tipologia di standard	aperto, *de facto*

Derivato da	—
Revisione	2015
Riferimenti	• Xiph.Org, *Vorbis I specification*, 2015 • xiph.org/vorbis
Conservazione	No
Racc. per la lettura	Specifico; raccomandato
Racc. per la scrittura	Nessuna raccomandazione

MusicXML	Formato di file
Nome completo	MusicXML™
Estensione/i	.musicxml
Specializzazione di	XML
Tipo MIME	application/vnd.recordare.musicxml
Sviluppato da	W3C Music Notation Community Group
Tipologia di standard	aperto, estendibile, *de facto*, testuale
Livello metadati	4; cfr. §2.8
Derivato da	MakeMusic® MusicXML, MuseData™, Humdrum™
Revisione	3.1 (2017)
Riferimenti	• W3C Community Group Final Report, *MusicXML 3.1*, 2017 • www.musicxml.com
Conservazione	Sì
Racc. per la lettura	Specifico; raccomandato per le partiture musicali
Racc. per la scrittura	Specifico; fortemente consigliato per partiture musicali

MIDI	Formato di file
Nome completo	Musical Instrument Digital Interface
Estensione/i	.mid, .midi
Magic number	MThd
Tipo MIME	audio/midi, application/x-midi
Sviluppato da	MIDI Manufacturers Association (MMA)
Tipologia di standard	aperto, estendibile, *de facto*, binario
Livello metadati	1
Derivato da	—
Revisione	1.1

Riferimenti	• www.midi.org/specifications • McGill, *Standard MIDI-File Format Spec. 1.1*. • McGill, *Standard MIDI Files ("SMF")*. • RFC-6295 • *General MIDI ("GM") Specifications*: GM1, Roland® *GS*, Yamaha® *XG*, GM2.
Conservazione	Sì, purché adottando solo strumenti General MIDI
Racc. per la lettura	Specifico; obbligatorio per la riproduzione di musica strumentale; raccomandato per partiture musicali
Racc. per la scrittura	Specifico; raccomandato per partiture musicali con orchestrazione rappresentabile da questo standard

2.9.1 Raccomandazioni per la produzione di documenti

1. 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: di conseguenza tutti i formati elencati –ad eccezione di AIFF– soddisfano tale caratteristica. Il wave, essendo in realtà un contenitore, è poi soggetto alla scelta del codec per rappresentare il contenuto audio. In questo caso si raccomandano codec aperti (come il "*raw*" non compresso, ovvero il PCM o il μ-law). La scelta di formati "*raw*," sebbene dettata spesso da esigenze contingenti, andrebbe evitata in quanto tali formati sono notoriamente muti (cfr. §1.2.2) e quindi può essere difficile ricostruire le caratteristiche, sia ambientali che tecniche, di registrazione e di campionamento.
2. Resta appannaggio delle singole organizzazioni valutare i parametri qualitativi in merito ai codec audio impiegati (ove sussistano come parametri variabili), tenendo conto ancora una volta di eventuali vincoli normativi, amministrativi e tecnologici. La scelta a priori dei parametri qualitativi può aiutare a standardizzare un processo e verificare che tutti gli organi coinvolti possano leggere o scrivere file con tali parametri.
3. Tra i formati sopra individuati, il WAV (con codifica PCM non compressa) e tutti i formati "*raw*" senza ausilio di compressione sono i formati più adatti alla conservazione. Nei casi in cui non si può evitare la compressione della forma d'onda, può essere adottato anche il formato FLAC.

2.10 Video

1. Un file (ovvero un flusso) video *strictu sensu* rappresenta una sequenza ordinata di immagini raster (dette fotogrammi — *frame* in inglese) riprodotte ad una velocità temporale fissa, espressa in numero di fotogrammi al secondo (alle volte si usa la sigla "fps" o, più impropriamente, l'unità di misura Hertz, "Hz"). Per il resto si applicano al video le stesse descrizioni relative alle immagini raster: risoluzione, profondità di colore, colorimetria, ecc.
2. Anche nel video, come per le immagini statiche e l'audio, è possibile implementare diversi algoritmi di compressione allo scopo di ridurre le dimensioni del file ovvero del flusso digitale. Quando la compressione riguarda individualmente e separatamente ogni singolo fotogramma si parla di compressione *intra*-frame (che è concettualmente identica a quella applicata alle immagini raster); quando la compressione utilizza le informazioni residue di un fotogramma (calcolate rispetto ai fotogrammi precedenti o antecedenti) si parla invece di compressione *inter*-frame.
3. Quest'ultima tipologia di compressione utilizza modelli matematici e statistici tipicamente più sofisticati di quelli usati nella prima tipologia, il che rende spesso l'operazione di codifica più lunga e laboriosa di quella di decodifica. Per il motivo di cui al punto precedente, la compressione inter-frame è raccomandata per applicazioni ove la tecnologia, le tempistiche e i costi si distribuiscano in modo da avere una generazione del video più lunga, laboriosa e in generale dispendiosa, a fronte di una minore complessità di decompressione a carico dei dispositivi riproduttori[33].
4. Al contrario, la compressione intra-frame, isolando ogni fotogramma in un'evidenza compressa a se stante, pur non fornendo in generale gli stessi rapporti di compressione di quella inter-frame, la rendono spesso –ma non sempre– preferibile in due particolari casi d'uso:

- quando sia frequente la riproduzione casuale del documento video, che consiste nel saltare da un punto all'altro della *timeline*[34] (discontinuità temporale) piuttosto che riprodurla sequenzialmente;
- quando il documento debba potersi modificare (cioè aggiungendo, inserendo, sostiutendo o eliminandone sue parti dalla timeline) senza alterarne la qualità.

5. Algoritmi di compressione intra-frame molto diffusi –non necessariamente associati a specifici formati di file– includono l'MJPEG, l'AVC-Intra (di proprietà di

[33] La compressione inter-frame favorisce dunque gli scenari d'uso ove la codifica (compressione) è effettuata una tantum da dispositivi centralizzati e computazionalmente potenti, a fronte di una riproduzione (decompressione) in serie, effettuata da una moltitudine di dispositivi riproduttori "*consumer*" con caratteristiche tecnologiche inferiori (inclusa la necessità di risparmio energetico). Un esempio tipico è la riproduzione cine-televisiva o in streaming on-demand, ove i riproduttori spaziano dai televisori ai *tablet*, agli *smartphone*.
[34] La timeline verrà definita in §2.12.

Panasonic®), la famiglia di codec proprietari ProRes® della Apple®, i vari video-codec ("VC") della SMPTE e il derivato proprietario dal codec VC-3: la famiglia DNxHD® della Avid®.
6. Algoritmi di compressione inter-frame largamente diffusi sono invece MPEG-2, MPEG-4, H.264 (e una sua variante proprietaria: AVCHD), XAVC (detto anche impropriamente "H.265") e la famiglia XDCAM™, anch'essa proprietaria e da non confondere con i formati del pacchetto di file usato nella medesima famiglia di prodotti (e in questo Allegato indicato in maniera tipograficamente diversa come "XDCAM", cfr. §2.12).

HEVC / H.265		CODEC
Nome completo	High Efficiency Video Coding	
Profili	Main, High; 13 livelli	
Codice FourCC	hevc, h265, x265	
Sviluppato da	Joint Collaborative Team on Video Coding (JCT-VC)	
Tipologia di standard	proprietario (*royalty* per varie tipologie d'uso), *de iure*	
Derivato da	MPEG-H Part 2	
Revisione	2018	
Riferimenti	• ISO/IEC 23008-2:2017 • ISO/IEC 23008-2/Amd-1:2018 • ITU-T Recommendation H.265 (2018) • ETSI TS-126-114 v15.7.0 (2019) • x265.org/hevc-h265	
Conservazione	No	
Racc. per la lettura	Generale; raccomandato	
Racc. per la scrittura	Generale; sconsigliato per archiviazione	

H.264 / AVC		CODEC
Nome completo	Advanced Video Coding	
Profili	21 profili e 20 livelli	
Codice FourCC	h264, x264, avc1, davc, vssh, v264	
Sviluppato da	Joint Collaboration Team on Video Codec (JCT-VC)	
Tipologia di standard	proprietario (*royalty* per uso commerciale), *de iure*	
Derivato da	MPEG-4 Part 2 (Visual), H.263	
Revisione	2014	
Riferimenti	• ISO/IEC 14496-10:2014 • ITU-T Recommendation H.264 (2019) • ETSI TS-126-114 v15.7.0 (2019) • x265.org/hevc-h265	
Conservazione	Sì, purché con profili e livelli almeno "High"	

Racc. per la lettura	Generale; obbligatoria la lettura
Racc. per la scrittura	Generale; raccomandato per archiviazione o conserva- zione di contenuti non cinetelevisivi ovvero master cinetelevisivi già sottoposti a elaborazione finale

MP4V	CODEC
Nome completo	MPEG-4 Part 2 MPEG-4 Visual
Profili	21 profili (tra cui SP, ASP, SSTP)
Codice FourCC	mp4v, mp42, avc1, v264
Sviluppato da	Video Coding Experts Group (VCEG), Moving Picture Experts Group (MPEG)
Tipologia di standard	proprietario (*royalty* per uso commerciale), *de iure*
Derivato da	H.263, MPEG-2 Part 2
Revisione	2014
Riferimenti	• ISO/IEC 14496-2:2004
Conservazione	No
Racc. per la lettura	Generale; obbligatoria la lettura
Racc. per la scrittura	Generale; raccomandato per archiviazione non a lungo termine; sconsigliato in conservazione ovvero quando sono necessari ulteriori metadati

H.263	CODEC
Nome completo	H.263 (XviD)
Profili	vari profili
Codice FourCC	h263, s263
Sviluppato da	Video Coding Experts Group
Tipologia di standard	proprietario (*royalty* per uso commerciale), *de iure*
Derivato da	MPEG-2 Part 2
Revisione	2006
Riferimenti	• ITU-T Recommendation Y.4414/H.263 (2005) • ETSI TS-126-141 v15.0.0 (2018)
Conservazione	No
Racc. per la lettura	Generale; obbligatoria la lettura
Racc. per la scrittura	Generale; sconsigliata la produzione

MPEG2		CODEC
Nome completo	MPEG-2 Part 2	
Profili	7 profili e 4 livelli	
Codice FourCC	mp2v	
Sviluppato da	Video Coding Experts Group, Moving Picture Experts Group	
Tipologia di standard	aperto, *de iure*	
Derivato da	MPEG-1	
Revisione	2018	
Riferimenti	• ISO/IEC 13818-2:2013	
Conservazione	Sì	
Racc. per la lettura	Generale; obbligatoria la lettura	
Racc. per la scrittura	Generale; sconsigliato per archiviazione e conserva-zione di contenuto cinetelevisivo professionale	

DNxHD		CODEC
Nome completo	DNxHD® / DNxHR™ [35]	
Profili	vari *bit-rate* in Mb/s / LB, SQ, HQ, 444, HQX	
Codice FourCC	AVdn	
Sviluppato da	Avid	
Tipologia di standard	proprietario (licenza commerciale), *de facto*	
Derivato da	SMPTE VC-3	
Revisione	Codecs LE 2.7.3 (2018)	
Riferimenti	• www.avid.com • avid.force.com Famiglia **ST2019** di standard/raccomandazioni SMPTE: • ST2019-1:2014	
Conservazione	No	
Racc. per la lettura	Speciale; consigliato in post-produzione cinetelevisiva	
Racc. per la scrittura	Speciale; sconsigliato per archiviazione/conservazione	

7. Il codec audio/video proprietario, indicato semplicemente come DNxHD®, dotato di compressione *lossy* e intra-frame, consiste in una specializzazione di alcuni profili del codec aperto VC-3 standardizzato dalla SMPTE (descritto altrove in questa sezione), adattato per l'utilizzo nella post-produzione non-lineare di documenti audiovisivi professionali. In particolar modo le essenze audio video sono vincolate a un insieme discreto di risoluzioni spaziali, profondità di colore, *framerate* e bit-rate costanti o variabili, ottimizzati per contenuti in risoluzione HD (1920×1080 punti,

[35] I nomi commerciali dei codec sono acronimi di '**D**igital **N**onlinear **E**xtensible **H**igh **D**efinition' e '**H**igh Dynamic **R**ange' rispettivamente.

si parla in questo caso di codec DNxHD®) e l'UltraHD 4K (3840×2160, si parla in questo caso di codec DNxHR™) che andranno successivamente montati e post-processati (ad esempio tramite operazioni di *color-grading* –in HDR o meno– *compositing*), possibilmente mediante operazioni "non-distruttive", cioè riducendo il più possible l'impatto che tali procedimenti di post-produzione avranno sulla qualità del documento masterizzato. La produzione del formato è chiusa e proprietaria, sebbene sia uno standard de facto del settore cinetelevisivo che diversi applicativi hardware e software sono in grado sia di produrre che di riprodurre. Nella loro declinazione nativa, le essenze audio/video con codec DNxHD® sono imbustate in un contenitore MXF (descritti nel §2.12) anch'esso dotato di particolari metadati. Tali metadati coadiuvano e semplificano il collegamento logico delle essenze audiovisive che, contenute in file separati, vengono associate insieme temporalmente da parte degli applicativi di montaggio non-lineare. È tuttavia possibile imbustare essenze DNxHD® n altri formati *wrapper*, come ad esempio il QuickTime.

ProRes		Codec
Nome completo	ProRes®	
Profili	Proxy, LT, 422, HQ, 4444, XQ, RAW	
Codice FourCC	apco, apcs, apcn, apch, ap4h, ap4x, aprh/aprn	
Sviluppato da	Apple Incorporated	
Tipologia di standard	proprietario (*royalty* per la produzione), *de facto*	
Derivato da	Apple Intermediate Codec	
Revisione	2018	
Riferimenti	• Apple, *ProRes White Paper* (2018) • Apple, *ProRes RAW White Paper* (2018)	
Conservazione	No	
Racc. per la lettura	Speciale; raccomandato nel montaggio cinetelevisivo	
Racc. per la scrittura	Sconsigliato per l'archiviazione o conservazione.	

8. L'Apple ProRes® è un codec audio/video proprietario dotato di compressione *lossy* e intra-frame, specializzato anch'esso per il montaggio non-lineare di documenti audiovisivi professionali. È dotato di diversi profili cui sono associate profondità di colore e bit-rate predeterminati (mentre non vi sono vincoli alla risoluzione o al *framerate* come accade per altri codec)

AV1		Codec
Nome completo	AOMedia Video 1	
Profili	3 profili e 10 livelli	
Codice FourCC	av01	
Sviluppato da	Alliance for Open Media	

Tipologia di standard	aperto, *de facto*
Derivato da	VP9, VP10
Revisione	2018
Riferimenti	• aomediacodec.github.io/av1-spec
Conservazione	No
Racc. per la lettura	Speciale; raccomandato per contenuti in *streaming* web
Racc. per la scrittura	Speciale; sconsigliato per archiviazione/conservazione

DASH / VP9 — CODEC

Nome completo	Video Partition structured video codec #9
Profili	–
Codice FourCC	VP90
Sviluppato da	Google
Tipologia di standard	proprietario (libero), *de facto*
Derivato da	VP8
Revisione	2014
Riferimenti	• Google, *VP9 bitstream & decoding process specification* v0.6 (2016) • ISO/IEC 14496-12:2015 • ISO/IEC 14496-12:2015 /Amd 2:2018 • Standard SMPTE ST2086:2014 • www.webmproject.org/vp9 • github.com/webmproject/vp9-dash
Conservazione	Sì
Racc. per la lettura	Speciale; raccomandato per contenuti in *streaming* web
Racc. per la scrittura	Speciale; sconsigliato per archiviazione/conservazione

CINEFORM — CODEC

Nome completo	SMPTE Video Codec #5
Profili	3 profili e 10 livelli
Codice FourCC	CFHD
Sviluppato da	Society of Motion Picture and Television Engineers
Tipologia di standard	aperto, *de iure*
Derivato da	GoPro® CineForm™
Revisione	2016
Riferimenti	Famiglia **ST2073** di standard SMPTE: • ST2073-1:2014, RP2073-2:2017, ST2073-3:2015, ST2073-4:2015, ST2073-5:2015, ST2073-6:2015 • gopro.github.io/cineform-sdk
Conservazione	No
Racc. per la lettura	Speciale; consigliato in post-produzione cinetelevisiva

Racc. per la scrittura	Speciale; sconsigliato per archiviazione/conservazione

VC-3		CODEC
Nome completo	SMPTE Video Codec #3	
Profili	–, HQ	
Codice FourCC	AVdn	
Sviluppato da	Society of Motion Picture and Television Engineers	
Tipologia di standard	aperto, *de iure*	
Derivato da	SMPTE VC-2 ("Dirac Pro")	
Revisione	2002	
Riferimenti	Famiglia **ST2019** di standard/prassi SMPTE: • ST2019-1:2014, RP2019-2:2016, ST2019-4:2014	
Conservazione	Sì	
Racc. per la lettura	Speciale; consigliato in post-produzione cinetelevisiva	
Racc. per la scrittura	Speciale; consigliato in post-produzione cinetelevisiva	

HDCAM		CODEC
Nome completo	HDCAM	
Profili	–, SR [Lite], SQ, HQ; diversi *bitrate*	
Codice FourCC	—	
Sviluppato da	Society of Motion Picture and Television Engineers	
Tipologia di standard	proprietario (libero), *de iure*	
Derivato da	Sony HDCAM™	
Revisione	2005	
Riferimenti	• Standard SMPTE ST367M:2001 *Type D-11 HDCAM picture compression and data stream format* • Standard SMPTE ST368:2002, *for digital television tape recording – 12.65mm Type D-11 format*	
Conservazione	No	
Racc. per la lettura	Speciale; consigliato per riprodurre e riversare repertori in nastri digitali HDCAM™ (formato D-11)	
Racc. per la scrittura	Speciale; sconsigliato per la produzione di nastri	

DIRAC Pro		CODEC
Nome completo	SMPTE Video Codec #2 "Dirac Pro"	
Profili	–, SR [Lite], SQ, HQ	

Codice FourCC	Bbcd
Sviluppato da	Society of Motion Picture and Television Engineers
Tipologia di standard	aperto, *de iure*
Derivato da	BBC "Dirac"
Revisione	2016
Riferimenti	Famiglie **ST2042** e **ST2047** di standard SMPTE: • ST2042-1:2090, RP2042-2:2009, RP2042-3:2010,
Conservazione	No
Racc. per la lettura	Speciale; consigliato in post-produzione cinetelevisiva
Racc. per la scrittura	Speciale; sconsigliato per archiviazione/conservazione

XAVC	CODEC
Nome completo	XAVC™
Profili	H.264 livello 5.2
Codice FourCC	Xavc
Sviluppato da	Sony
Tipologia di standard	proprietario (*royalty* per uso commerciale), *de facto*
Derivato da	MPEG-2, H.263, MPEG-4 Part 2
Revisione	2018
Riferimenti	• www.xavc-info.org, *XAVC profiles and Operating Points* v1.20 (2016)
Conservazione	No
Racc. per la lettura	Speciale; consigliato in post-produzione cinetelevisiva
Racc. per la scrittura	Da *non* utilizzare per la produzione di documenti.

XDCAM	CODEC
Nome completo	XDCAM™
Profili	DVCAM, IMX, SD, EX, HD422, HD, SD422
Codice FourCC	xdv1, xdv2, …, xdv9, xdva
Sviluppato da	Sony
Tipologia di standard	proprietario (*royalty* per uso commerciale), *de facto*
Derivato da	MPEG-4 Part 2, MPEG-2 Part 2, DV
Revisione	2018
Riferimenti	• sonybiz.net
Conservazione	No
Racc. per la lettura	Speciale; consigliato in post-produzione cinetelevisiva
Racc. per la scrittura	Speciale; sconsigliato per archiviazione/conservazione

BRAW	FORMATO DI FILE
Nome completo	Blackmagic RAW
Profili	XXXXX
Codice FourCC	Braw
Sviluppato da	Blackmagic Design
Tipologia di standard	proprietario
Derivato da	CinemaDNG™
Revisione	2018
Riferimenti	• www.blackmagicdesign.com
Conservazione	No
Racc. per la lettura	Speciale; consigliato in post-produzione cinetelevisiva
Racc. per la scrittura	Speciale; sconsigliata la produzione

AVC-INTRA	CODEC
Nome completo	AVC-Intra™
Profili	2 *bit-rate* e 6 risoluzioni
Codice FourCC	ai[5|1][p|q|2|3|5|6]
Sviluppato da	Sony
Tipologia di standard	proprietario (*royalty* per uso commerciale), *de iure*
Derivato da	MPEG-4 Part 10
Revisione	2014
Riferimenti	• prassi raccomandata SMPTE RP2027:2012 • ISO/IEC 14496-10:2014
Conservazione	No
Racc. per la lettura	Speciale; raccomandata la lettura
Racc. per la scrittura	Da *non* utilizzare per la produzione di documenti.

9. AVC-Intra deriva dal profilo High 10 Intra di H.264.

MJPEG	FORMATO DI FILE / CODEC
Nome completo	Motion JPEG
Profili	A, B
Codice FourCC	mjpg, mjpa, mjpb
Sviluppato da	comunità open source
Tipologia di standard	aperto, *de facto*
Derivato da	JPEG
Revisione	1998
Riferimenti	• RFC-2435
Conservazione	Sì
Racc. per la lettura	Generico con obbligo in lettura

Racc. per la scrittura	Generico; consigliato per contenuti non cinetelevisivi che non necessitano di ulteriori metadati

10. Il formato MJPEG, come recita il suo stesso acronimo, «Motion JPEG», è un formato video ove l'essenza video è codificata come una sequenza di immagini raster nel formato JPEG, adottando dunque una compressione *lossy* e intra-frame. Nonostante il formato sia datato, la sua semplicità di codifica lo rende ancora raccomandato in tutti gli utilizzi non professionali ove il documento informatico non abbia particolari requisiti di qualità (o comunque requisiti di qualità non superiore a quella della compressione JPEG scelta per l'essenza MJPEG).

2.10.1 Raccomandazioni per la produzione di documenti

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. Pe questo motivo le scelte dei formati avranno un'elevata variabilità in base alla finalità d'uso e ai vincoli tecnologici ad esse collegati.[36]
2. Lo strumento della valutazione di interoperabilità (cfr. §3.1) 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 – cioè avulse da casi d'uso peculiari di specifici settori quali la produzione, postproduzione e conservazione dei contenuti audiovisivi– sono quelli aperti e standard de iure, quali:

[36] Ad esempio, un'evidenza video potrebbe essere usata da un certo numero di utenze poco dopo la sua formazione, che richiedono tecnologie di elaborazione in tempo reale, a fronte delle quali è necessaria la produzione o la transcodifica in un codec ad elevato rapporto di compressione. Potrebbe anche esserci un vincolo di conservazione dell'evidenza originale piuttosto che di quella elaborata, che potrebbe avere o meno diversi requisiti di qualità. A fronte di questa seconda finalità, potrebbe essere pensato un procedimento per cui l'evidenza originale viene formata impiegando un codec senza compressione e la massima risoluzione possibile, mentre il file elaborato segue un ciclo vita diverso, al termine del quale possa o debba essere distrutto. Qualora anche il file elaborato vada conservato, potrebbe essere necessario prevedere un solo flusso, a fronte del quale andrebbe effettuata una scelta diversa riguardo alla transcodifica intermedia, allo scopo di assolvere ai vincoli qualitativi iniziali.

- 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.
- Qualunque codec senza compressione, ovvero con una compressione priva di perdite. In tali casi sono valutabili tutte le rappresentazioni del documento video mediante sequenze di file (cfr. §2.12), ove i singoli file adottano formati con i medesimi requisiti di interoperabilità (cfr. §2.6) inclusa l'assenza di compressione ovvero la compressione *lossless*.

4. Nel caso di esigenze specifiche, si possono valutare anche altri codec, quali ad esempio:

- VP9 (anche detto "DASH") qualora il contenuto sia stato generato per applicazioni multipiattaforma (prevalentemente online) e non sia soggetto ad ulteriori modifiche.
- VC-3 (in particolar modo la sua variante "commerciale": l'Avid® DNxHD®) qualora il contenuto sia destinato al montaggio video (incluse successive archiviazioni del girato originale) e in assenza di ulteriori vincoli tecnologici.

5. Resta appannaggio delle singole organizzazioni valutare i parametri qualitativi in merito ai codec audio impiegati (ove sussistano come parametri variabili), tenendo conto ancora una volta di eventuali vincoli normativi, amministrativi e tecnologici. La scelta a priori dei parametri qualitativi può aiutare a standardizzare un processo e verificare che tutti gli organi coinvolti possano leggere o scrivere file con tali parametri.
6. I codec video più adatti alla conservazione –comunque tutti dotati di compressione con perdita– sono l'MPEG-4 Part-10

2.11 Sottotitoli, didascalie e dialoghi

1. Astrattamente, esiste una terza rappresentazione del suono legata ai *dialoghi* (rispetto a quelle descritte in §2.9) che consiste nella trascrizione più o meno letterale di ciò che una o più voci esprimono in una linea temporale, più eventualmente altre informazioni di contesto e ambiente.

2. La necessità di rappresentare il parlato in una forma leggibile è prevalentemente legata a quattro casi d'uso, in alcuni casi concorrenti fra loro:

a) conservare dialoghi in forma scritta, eliminando la complessità tecnica o giuridica di conservare una registrazione multimediale degli stessi;
b) arricchire un dialogo con ulteriori informazioni quali, ad esempio, il nome (o la qualifica) dei partecipanti, il momento esatto e la *consecutio* temporale con cui le frasi sono pronunciate;
c) complementare il dialogo con informazioni relative all'ambiente o al contesto, a beneficio di spettatori ipoudenti o non-udenti, come nel caso di riproduzioni audiovisive;
d) fornire la traduzione di testi e dialoghi in lingue diverse da quella o quelle determinate preventivamente per i destinatari individuati per un particolare contenuto audiovisivo.

3. Le trascrizioni giuridiche ottemperano complessivamente ai casi d'uso 1 e 2 (e 4 in caso di testimoni stranieri, ove siano effettuate da traduttori giurati). Invece le trascrizioni di dialoghi in ambito cinetelevisivo –chiamate in inglese, nella loro accezione più generale, *timed-text*– si dividono tradizionalmente in tre categorie:

- sottotitoli (in inglese *subtitles*, abbreviati in *"subs"*) — ottemperano al solo caso d'uso 4;
- sottotitoli per non-udenti (in inglese *deaf or hard-of-hearing*, ovvero DHH) — ottemperano al solo caso d'uso 3;
- didascalie (in inglese *closed captions*, ovvero CC) — ottemperano al caso d'uso 2 sebbene, sovente combinati con tipologie di sottotitoli di cui sopra, ottemperino anche ai casi 3 o 4.

4. Il settore dell'intrattenimento sta venendo rivoluzionato dalla comparsa di innovativi servizi di fruizione tramite internet dei contenuti audiovisivi sotto forma di flussi multimediali (i cosiddetti servizi SVOD, cioè *streaming video on-demand*); tali servizi consentono, tra le altre cose, una *localizzazione*[37] molto più puntuale, precisa e multilingua, che li rende ancora più pervasivi e globali.
5. La normativa di riferimento per sottotitoli e didascalie dal punto di vista di accessibilità e qualità (soprattutto quando sono mostrate in sincronia e sovraimpressione, rispettivamente, con essenze audio e video) è regolamentata in particolare dalla norma ISO/IEC 20071-23 del 2018.

TTML	FORMATO DI FILE
Nome completo Timed Text Markup Language	

[37] La localizzazione è un processo riadattamento incrociato (in inglese *versioning*) di un contenuto rispetto a molteplici territori. Quando ciò riguardi il solo adattamento linguistico (anziché comprendere anche aspetti culturali, religiosi, giuridici) di audiovisivi, si parla più semplicemente di ri*edizione*, che comprende il video, il *doppiaggio* per l'audio e la sottotitolazione (trascrizione di dialoghi e didascalie).

Estensione/i	.ttml, .dfxp
Specializzazione di	XML (*namespace* tt, tt[p|s|m] e altri)
Tipo MIME	application/ttml+xml
Sviluppato da	World Wide Web Consortium Society of Motion Picture and Television Engineers
Tipologia di standard	aperto, estendibile, retrocompatibile, *de iure*, testuale
Livello metadati	4
Derivato da	Distribution Format Exchange Profile (DFXP)
Revisione	1.0 (2018)
Riferimenti	• W3C Recommendation *TTML 2*, 2018 • W3C Recommendation *TTML 1 (3rd Ed.)*, 2018 • ETSI EN-303-560 v1.1.1 (2018) Famiglia **2052** di standard/raccomandazioni SMPTE: • OV2052-0:2013 • ST2052-1:2013 • ISO/IEC 14496-30:2018 • EBU – Tech 3380, *EBU-TT-D subtitling distribution format* v1.0.1 (2018)
Conservazione	Sì; cfr. §2.8
Racc. per la lettura	Speciale; obbligatorio in campo cinetelevisivo
Racc. per la scrittura	Speciale; raccomandato in campo cinetelevisivo

6. Il formato TTML (precedentemente chiamato DFXP) è, ad oggi, la declinazione più generica e completa relativamente alle possibilità tecniche e operative di aggiungere ad un documento audiovisivo una o più trascrizioni differenti. Il formato, adottato sempre di più come standard a livello europeo (*de iure*: cfr. norma EN-303-560 della ETSI) e internazionale[38] supporta qualunque tipo di lingua scritta e sue eventuali caratteristiche diacritiche (, integrando informazioni sulla presentazione dei sottotitoli all'interno del riquadro del video (posizione,[39] colore, dimensione, parametri tipografici completi), relativamente all'asse temporale (apparizione, scomparsa e altri effetti di animazione).

IMSC1	FORMATO DI FILE
Nome completo	Internet Media Subtitles and Captions
Estensione/i	.ttml

[38] Esistono diversi standard *de facto*, quali ad esempio le Linee guida di NETFLIX sulla sottotitolazione (in inglese).
[39] È anche possibile specificare la profondità ("*Z-depth*") rispetto alla parallasse nel caso di sottotitoli su contenuti stereoscopici (S3D).

Specializzazione di	TTML
Tipo MIME	application/ttml+xml
Sviluppato da	World Wide Web Consortium
Tipologia di standard	aperto, estendibile, retrocompatibile, *de iure*, testuale
Livello metadati	4
Derivato da	W3C Timed Text Markup Language, versione 2
Revisione	1.1 (2018)
Riferimenti	• W3C Recommendation, *TTML profiles for IMSC 1.1*, 2018 • Netflix, *What does a properly formatted TTML file look like?* • Netflix, *Italian Timed Text Style Guide*, 2018
Conservazione	Sì; cfr. §2.8
Racc. per la lettura	Speciale; raccomandato in campo cinetelevisivo
Racc. per la scrittura	Speciale; raccomandato in campo cinetelevisivo

7. Il W3C ha deciso di standardizzare ulteriormente il generico formato TTML, derivando perciò le specifiche IMSC1 adottate, tra le altre cose dal formato di master interoperabile (IMF, cfr. §2.12). IMSC1 ha due profili: uno ove i sottotitoli sono rappresentati mediante solo testo, l'altro ove sono sottotitolo è rappresentato da un'immagine (tipicamente nel formato PNG con supporto completo del canale alfa, cfr. §2.6). L'esigenza di introduzione di questo profilo è stata necessaria per il diffondersi di notevoli dialetti di TTML da parte di diversi organi di standardizzazione, che hanno portato il formato originale –originariamente disegnato per essere il più interoperabile possibile– a creare notevoli problemi di compatibilità.[40] Il profilo IMSC1, oltretutto, è stato concepito anche per ottemperare ad esigenze proprie della distribuzione dei contenuti, quali l'adattamento dei dialoghi a diverse revisioni del contenuto audiovisivo, alla riedizione dei dialoghi in altre lingue o secondo esigenze diverse (cfr. distinzioni fra le varie tipologie di dialoghi sincronizzati all'inizio della sezione), alla transcodifica dei sottotitoli in altro formato.

EBU-TT	FORMATO DI FILE
Nome completo	EBU Timed Text (EBU-TT)
Estensione/i	.xml
Specializzazione di	XML (*namespace* ebutt[m|s|dt])

[40] Tale problema si è aggravato per il fatto che –come accade anche per molti altri formati di file per sottotitoli e didascalie– vi è un elevato riutilizzo della medesima estensione di file per formati anche molto diversi fra loro, cfr. §1.1.3.

Tipo MIME	application/ttml+xml
Sviluppato da	European Broadcasting Union
Tipologia di standard	aperto, estendibile, retrocompatibile, *de iure*, testuale
Livello metadati	2
Derivato da	W3C Timed Text Markup Language, versione 1
Revisione	1998
Riferimenti	Famiglia **33x0** di standard tecnici EBU (2017): • Tech 3350, *part 1: subtitling format* v1.2 • Tech 3360, *part 2: mapping EBU STL to EBU-TT* v1.0 • Tech 3370, *part 3: mapping EBU STL to EBU-TT* v1.0 • Tech 3380, *subtitling distribution format* v1.0.1 (2018) • Tech 3390, *part M: metadata definitions* v1.0
Conservazione	No; cfr. §2.8
Racc. per la lettura	Speciale; obbligatorio in campo cinetelevisivo
Racc. per la scrittura	Speciale; raccomandato il campo cinetelevisivo ove richiesto da applicativi o capitolati

8. La EBU ha adattato il formato TTML agli specifici vincoli del mondo dell'audiovisivo, definendo così una specifica del medesimo formato, chiamato **EBU-TT**.
Analogamente, anche la SMPTE, ha definito una specifica TTML per i medesimi scopi, chiamata **SMPTE-TT**, che però è stata adottata prevalentemente dalla versione dei pacchetti per il cinema digitale (DCP) e quindi raccomandata in altra sezione di questo Allegato ad essi dedicata, cfr. §2.12.

STL	FORMATO DI FILE
Nome completo	EBU Subtitling Data Exchange Format
Estensione/i	.STL
Magic number	0x3?3?3? STL
Tipo MIME	–
Sviluppato da	European Broadcasting Union
Tipologia di standard	aperto, estendibile, *de iure*, binario
Livello metadati	1
Derivato da	–
Revisione	1.0 (1991)
Riferimenti	• EBU – Tech 3264 (1991)

Conservazione	No
Racc. per la lettura	Speciale; raccomandato in campo cinetelevisivo
Racc. per la scrittura	Speciale; sconsigliato per la produzione e archiviazione

9. Il formato STL (da non confondersi con l'omonimo formato per modelli 3D per applicazioni stereolitografiche, cfr. §2.7) è stato per anni lo standard di riferimento in campo cinetelevisivo, perciò un grandissimo numero di contenuti risultano dotati di sottotitoli e didascalie in questo formato, in svariate lingue. Il formato tuttavia contiene solo indicazioni relative ai TimeCode ove i sottotitoli appaiono e scompaiono, al testo della lingua e a pochissime indicazioni visive su come presentare i sottotitoli, peraltro ampiamente non sempre utilizzate o rispettate dagli applicativi rispettivamente di creazione e riproduzione video.

2.11.1 Raccomandazioni per la produzione di documenti

1. Per la produzione di sottotitoli e didascalie in ambito cinetelevisivo, 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.

2.12 Contenitori e pacchetti di file multimediali

1. La registrazione e riproduzione di immagini e suoni contemporaneamente necessita innanzi tutto del sincronismo tra gli elementi costituenti un flusso multimediale. Per questo motivo tali flussi o file (a seconda del livello di astrazione adattato) sono disposti in una "linea temporale virtuale" (*timeline*, in inglese) che adotta un riferimento temporale autonomo[41] chiamato TimeCode.
2. Un esempio di sintassi per un TimeCode sincronizzato con un flusso video digitale (la cui unità indivisibile è il fotogramma) è, ad esempio, "**04:23:56:**07"[42]. Un esempio di sintassi per un TimeCode sincronizzato con un flusso audio digitale

[41] Per analogia, mentre la datazione di un'evidenza informatica mediante marcatura temporale è un riferimento temporale "assoluto", la temporizzazione di uno specifico punto della timeline è un riferimento temporale "relativo", che diviene assoluto solo nel momento in cui la riproduzione della timeline inizia ad un determinato istante temporale.
[42] Leggasi «4 ore, 23 minuti, 56 secondi e 7 fotogrammi» (i fotogrammi si azzerano al secondo in base al numero di fps, cfr. §2.10).

(la cui unità indivisibile è il campione sonoro) è, ad esempio, "**00:06:32.**01436"[43]. Più in generale i TimeCode sono normati da standard di riferimento, quali la raccomandazione BT.1366 della ITU, la famiglia di standard **ST12** della SMPTE, la EBU *Tech 3097*.

3. Nel caso multimediale un flusso o file contiene spesso evidenze informatiche diverse per audio, video, e altri tipi di metadati (quali ad esempio sottotitoli, informazioni per i non udenti, uno o più TimeCode di riferimento, ecc.), che sono complessivamente chiamate essenze. Esse –ciascuna potenzialmente codificata con parametri e algoritmi diversi– sono racchiuse in un unico file contenitore, il quale ne facilita il sincronismo (eventualmente grazie alla presenza di una o più essenze TimeCode) e le descrive allo scopo di migliorarne la riproduzione o l'utilizzo.

4. Le essenze possono a loro volta contenere più *canali*. Nel caso dell'audio essi sono associati ad un particolare dispositivo di riproduzione ovvero alla direzione spaziale della sorgente sonora, p.es. canale dei totali sinistra (L_t), canale centrale (C), canale degli effetti a bassa frequenza (LFE), canale per il surround di destra (R_s). Nel caso di immagini statiche o video, i canali possono codificare un singolo canale cromatico dello spazio-colore riprodotto (cfr. §2.6), ad es. rosso (R), verde (G), blu (B) ovvero uno di due stereogrammi (occhio destro e occhio sinistro):

5. Così come una partitura musicale descrive per prima cosa tutti gli strumenti coinvolti e contiene le informazioni –sia comuni che specifiche per ogni parte– affinché un'orchestra possa rieseguire il brano, allo stesso modo un contenitore multimediale può contenere uno o più dei seguenti fattori (la cui presenza possibile e/o obbligatoria dipende dallo specifico formato della busta, cfr. §1.1.1):

- informazioni globali circa il documento multimediale (e.g. titolo del film/brano, nome dell'autore/regista, anno di produzione, durata nominale, livello di censura, data/ora di masterizzazione del file, ecc.);
- numero e tipologia delle essenze presenti e loro suddivisione in "tracce" per la riproduzione; per ciascuna essenza possono essere presenti altri metadati come, ad esempio:
    - nome o altro tipo di descrizione dell'essenza;
    - codec dell'essenza (ad esempio tramite codice FourCC);
    - velocità dell'essenza video (fotogrammi/secondo) o audio (campioni/secondo);
    - nome e numero di canali dell'essenza video (e.g. RGB, RGBα, CMYK, …; occhio destro/sinistro) o audio (e.g. monoaurale, stereo, quadrifonico, '5.1', '7.1', Dolby-E®, '22.1', Dolby® Atmos™, …);
    - in caso di essenza multi-canale, eventuale nome dei singoli canali;

[43] Leggasi «6 minuti, 32 secondi e 1436 campioni» (i campioni si azzerano al secondo in base al sample rate: p.es. 44100Hz, cfr. §2.9).

    - profondità digitale dell'essenza immagine/video (bit/pixel) o audio (bit/campione);
    - *data-rate* minimo, medio e/o massimo di ciascun essenza.
    - lingua utilizzata per il parlato nell'audio (se presente) ovvero nei sottotitoli;
- 1 essenza video (nel caso di video stereoscopico potrebbe essere presente un'unica essenza con due canali, ovvero due essenze video distinte — una per ciascuno stereogramma, associati rispettivamente all'occhio destro e sinistro;
- 3 essenze audio (e.g. 'suoni ed effetti', 'colonna sonora italiana', 'colonna sonora inglese');
- 5 essenze sottotitoli (e.g. 'sottotitoli in italiano', 'sottotitoli in italiano per non-udenti', 'didascalie in inglese, 'sottotitoli in francese', 'didascalie in tedesco per non-udenti').
- 1 essenza TimeCode[44].

6. Come primo vantaggio, l'utilizzo di un formato contenitore consente ad applicazioni di conservazione e archiviazione di estrapolare molte informazioni da un file complesso e di grandi dimensioni (quale spesso è il file multimediale) senza dover implementare tutti i possibili codec necessari per decodificarne le essenze audio\video (cosa necessaria, invece, per poter riprodurre il file). Inoltre, il formato contenitore separa molto nettamente, al suo interno, le evidenze informatiche delle singole essenze, permettendone un'estrapolazione segmentata, molto più efficace e spesso anche con economia di spazio di archiviazione e tempi di elaborazione.
7. In alcuni casi di pertinenza multimediale, ad esempio, una o più essenze vengono ricodificate allo scopo di alternarne il *data-rate* o semplicemente per cambiare il codec utilizzato: si parla in tal caso, di *transcoding* delle essenze. Quando si cambia una o più essenze per un file multimediale (ad esempio per togliere una colonna sonora non più necessaria, aggiungere una traccia sottotitoli, ovvero sostituire una traccia video monoscopica con una stereoscopica) si parla di *re-wrapping*. Quando infine si cambia il formato di busta contenente l'intero file multimediale, ma mantenendo invariate le essenze, si parla di *trans-wrapping* — che è tipicamente un'operazione molto meno computazionalmente onerosa.

[44] Un esempio in cui sono presenti più tracce TimeCode è quando sono necessari più riferimenti temporali diversi (allo scopo di sincronizzare il flusso con altri sistemi informativi o altri flussi): ad esempio potrebbe essere presente un TimeCode "assoluto" espresso in orario diurno sulle 24 ore e uno "relativo" a un inizio temporale arbitrario rappresentato da "**00:00:00:**00".

Figura 1 Esempio di una timeline semplice di un file multimediale con una traccia video (monoscopica) e una traccia audio biaurale (2 canali stereo).

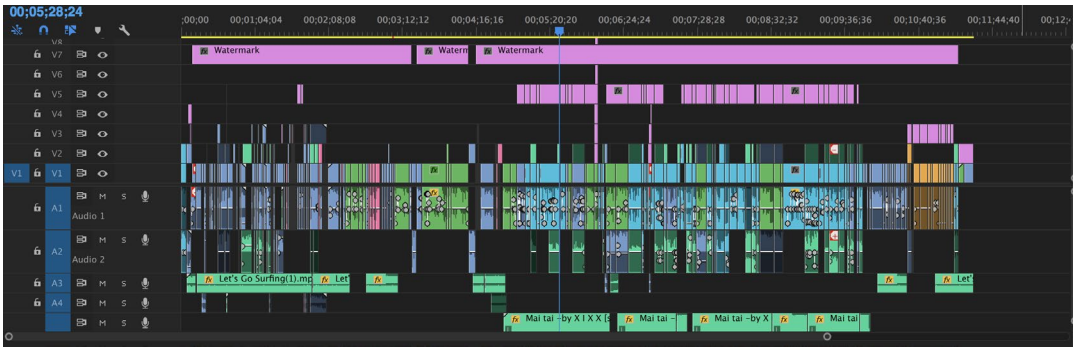

Figura 2 Esempio di una timeline di un file multimediale più complesso, con 9 tracce video (di cui una stereoscopica), 18 tracce audio mono o biaurali (cioè a 1 o 2 canali), 1 traccia TimeCode e diverse tracce sottotitoli nascoste.

8. Un'alternativa all'impiego di buste multimediali (cfr. §1.1.1) come sopra descritta è costituita dall'organizzazione delle singole essenze in file separati, poi riorganizzati in un pacchetto di file, ove le essenze siano logicamente associate fra loro in vari modi, come definito nel §1.1.2. Esempi di pacchetti multimediali sono ad esempio il pacchetto di master interoperabile (IMP), il pacchetto per il cinema digitale (DCP) e il pacchetto XDCAM, cfr. §2.12.
9. Per quanto riguarda le essenze video –nel caso di semilavorati o master di elevata qualità– è spesso preferito l'impiego della sola *naming convention*, ove ogni fotogramma del video è contenuto in un file separato, mentre i file sono numerati in sequenza cronologica (rispetto alla loro timeline relativa), costituendo dunque un pacchetto chiamato "sequenza di fotogrammi" (*frame sequence*, ovvero *frame-per-file* in inglese). Formati di immagini raster spesso adottati per tali sequenze di fotogrammi sono l'OpenEXR, il DPX (§2.6), così come il TIFF (soprattutto nei pacchetti DCDM), e il CinemaDNG (§2.12).

MXF		FORMATO CONTENITORE
Nome completo	Material Exchange Format	
Estensione/i	.mxf	
Magic number	–	
Tipo MIME	application/mxf	
Sviluppato da	Society of Motion Picture and Television Engineers	
Tipologia di standard	aperto, estendibile, *de iure*, binario	
Livello metadati	4	
Derivato da	–	
Revisione	2018	
Riferimenti	• RFC-4539 (2006) Documenti SMPTE, prevalentemente standard ('ST'): • base: ST377; EG41, EG42 • schemi operativi: ST378, ST390–393, ST407, ST408 • contenitori: ST379, ST381–389, ST394, ST405 • registri e dizionari di metadati: ST380, ST436; RP210, RP224 • www.smpte.org	
Conservazione	Sì, solo usando codec adatti alla conservazione	
Racc. per la lettura	Speciale; obbligo in lettura in tutti i settori audiovisivi	
Racc. per la scrittura	Speciale; raccomandato per la produzione, lo scambio e l'archiviazione di contenuti cinetelevisivi	

10. Il formato contenitore MXF è stato introdotto dalla SMPTE in un tentativo di standardizzare un formato per contenuti e flussi (cfr. §1.1.1) audiovisivi professionali che fosse utilizzabile lungo l'intera filiera cinetelevisiva, supportando il maggior numero di procedimenti: dalla produzione dei contenuti (camere digitali e animazione), attraverso tutte le fasi della post-produzione, sino alla distribuzione su diverse piattaforme (cinema digitale, televisione digitale terrestre e satellitare, streaming via internet, ecc.) e successiva archiviazione. Il formato MXF è regolamentato da diversi standard dell'SMPTE (fare riferimento alla tabella qui sopra), ma il principale fra loro è l'ST377. Il contenitore è professionale in quanto estremamente flessibile nell'imbustare più essenze di tipologie differenti, e supportare un'ampia gamma di metadati interni, associabili a diversi livelli semantici del contenuto, incluse specifiche posizioni della timeline)[45]. I metadati supportati sono estendibili: oltre ad una serie di metadati minimi supportati, la SMPTE mantiene

[45] Metadati associabili a particolari posizioni della *timeline* sono, ad esempio, commenti e annotazioni di montatori, coloristi, tecnici digitali o assistenti alla regia; l'elenco delle persone reali (e.g. ospiti, attori) ovvero finzionali (personaggi) inquadrati in quel momento.

un registro ove ad ogni semantica è associato un codice binario (UL, ovvero *unified label* in inglese). Il registro è consultabile applicativamente o interattivamente sul sito smpte-ra.org, ovvero direttamente sotto forma di XML (§2.2). Per supportare molteplici casi di utilizzo, procedure di elaborazione e di tecnologie di stoccaggio e trasporto dei contenuti e dei flussi multimediali, le essenze stesse possono essere disposte in una molteplicità di strutture differenti all'interno dell'evidenza informatica complessiva. Ad esempio, la registrazione simultanea e in tempo reale, nel medesimo file MXF, di due flussi video provenienti da una cinepresa stereoscopica e di flussi audio provenienti da più microfoni disposti in scena, prevede che le essenze audio e video siano disposte in maniera "interlacciata": il contenitore MXF avvolge perciò, sequenzialmente, ciascun fotogramma, seguito dalle essenze audio relative alla durata del medesimo fotogramma, seguito dal fotogramma successivo e così via.

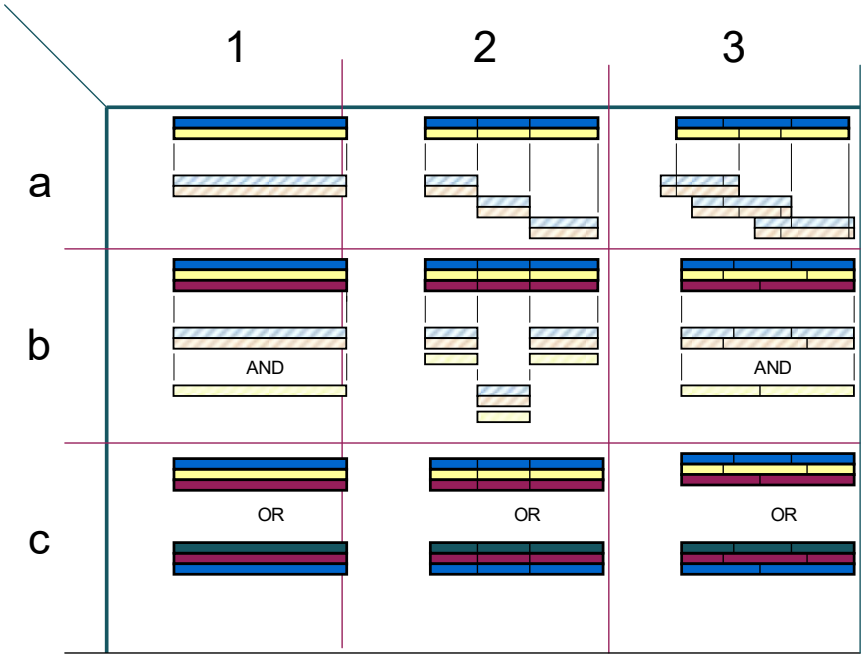

Figura 3 Tassonomia degli schemi operativi del formato MXF: sulle ascisse è rappresentata la complessità della timeline (indicata con un numero ordinale); sulle ordinate quella dei pacchetti (indicata con una lettera minuscola).

11. Per completare la descrizione di insieme del formato MXF, la SMPTE fornisce degli schemi operativi (*operational patterns* in inglese, abbreviati in **OP**), per i quali si invita a far riferimento ai relativi standard in tabella (oltre che allo standard principale ST377). In questi schemi operativi una timeline è astratta ulteriormente separando la suddivisione (e la "complessità") delle sue tracce e canali, dall'insieme di file ("pacchetti materiali") da cui è logicamente composta. In una tale tassonomia un contenuto multimediale (un'unica timeline) è rappresentato da un *pacchetto di contenitori MXF*, ciascuno contenente una o più essenze. La costruzione della

timeline logica del contenuto a partire dai file MXF costituenti può avvenire in modi diversi, la cui complessità dipende da una duplice organizzazione virtuale. Da una parte vi è quella delle essenze lungo la timeline (cfr. l'elenco alfabetico per righe, di complessità crescente, in **Figura 2**):

a. ogni istante della timeline è riferito a una o più essenze contenute in un unico file;
b. ogni istante della timeline è riferito a più essenze contenute in uno o più file;[46]
c. vi sono due o più timeline alternative[47], ciascuna delle quali è riferita a una o più essenze contenute in uno o più file;

Dall'altra vi è l'organizzazione di come le varie essenze della timeline siano distribuite in un pacchetto di uno o più file (cfr. l'elenco numerico per colonne, di complessità crescente, nella medesima **Figura 2**):

1. il pacchetto contiene essenze usate interamente e della medesima durata della timeline;[48]
2. il pacchetto contiene più sequenze di essenze, usante interamente, adiacenti fra loro lungo la timeline;[49]
3. il pacchetto contiene più sequenze di essenze, ciascuna delle quali (contenuta in file differenti) può essere usata parzialmente o interamente.[50]

12. Ad esempio, lo schema operativo **OP**1a corrisponde a quello della singola clip audio\video con le cui essenze sono tutte di pari durata[51], dove cioè un singolo file contiene all'interno l'audio, il video e gli eventuali sottotitoli e può quindi essere riprodotto autonomamente dall'inizio alla fine. Lo schema operativo **OP**3c, invece, prevede un montaggio asincrono tra le varie tracce, con i contenuti potenzialmente estratti da essenze di durata più lunga e contenuta in diversi file del pacchetto e, in più, la possibilità che vi siano più montaggi alternativi tra cui scegliere. Gli schemi operativi permettono di indicare queste differenti tipologie mediante metadati nei file MXF costituenti il pacchetto, ma anche di vincolarli logicamente fra di loro, cosa particolarmente utile sia nel caso in cui i file perdano l'affinità per referenza

[46] Nel gergo del montaggio audiovisivo, si dice che tali essenze sincronizzate sono 'in *gang*' (letteralmente, "raggruppate") fra loro.
[47] La selezione di una delle timeline alternative, o il passaggio dall'una all'altra, non fanno parte dello schema operativo, anche se ciò è solitamente computo in automatico dall'applicazione che riproduce il contenuto, ovvero da una scelta manuale dello "spettatore".
[48] In questo caso la timeline è fatta da una sola sequenza (o "taglio"), la cui riproduzione coinvolge per intero le essenze di uno o più file.
[49] In questo caso la timeline è fatta da più sequenze (o "tagli"), fra loro contigue nel tempo, come avviene nel caso d'uso della *playlist*.
[50] Questo è lo scenario tipico usato nel montaggio audiovisivo, ove porzioni di clip "sorgenti" sono montate in una timeline, tipicamente separate da tagli netti.
[51] Un altro schema operativo non descritto in Figura è l'**OP**-Atom: una variante dell'**OP**1a ove le essenze sono inoltre tutte auto-contenute in un singolo file.

(venendo archiviati in percorsi differenti o non stabili nel tempo), sia nel caso di successive modificazioni del contenuto. In quest'ultimo caso, modifiche parziali del contenuto multimediale vengono effettuate sui soli file MXF che contengono le essenze interessate, senza ricostruire o modificare il contenuto nella sua interezza. Più avanti sono descritti alcuni pacchetti di file che implementano proprio questa flessibilità degli schemi operativi del contenitore MXF: il formato di master interoperabile (IMF) e il pacchetto di cinema digitale (DCP).

13. Per tutti i motivi sopra esposti il contenitore MXF è fortemente raccomandato come formato contenitore d'elezione per la produzione e il riversamento di qualunque contenuto multimediale (audio, video, dialoghi), accompagnato ovviamente da una scelta adeguata dei codec da impiegare.

MP4	FORMATO CONTENITORE
Nome completo	MPEG-4, Part 14
Estensione/i	.mp4, .m4a, .m4v
Magic number	0x00000000 ftyp
Tipo MIME	video/mp4, audio/mp4
Sviluppato da	Society of Motion Picture and Television Engineers
Tipologia di standard	aperto, estendibile, retrocompatibile, *de iure*, binario
Livello metadati	3
Derivato da	Apple® QuickTime™, MPEG-4 Part 12
Revisione	2018
Riferimenti	• ISO/IEC 14496-14:2018 • ISO/IEC 14496-12:2015 • RFC-4337 (2006) • mp4ra.org
Conservazione	Sì
Racc. per la lettura	Generico; obbligatorio
Racc. per la scrittura	Generico; altamente consigliato per la produzione audiovisiva generica di contenuti auto-consistenti

14. La parte 14 del gigantesco standard MPEG-4 della ISO, ma anche contenuta nella RFC-4337 (più colloquialmente chiamata *MP4*) descrive uno standard *de iure* derivato moltissime sue caratteristiche dal formato Apple QuickTime (si legga sotto). Per questo motivo, e anche a causa della sua elevata versatilità sia nel numero e tipologie di essenze supportate al suo interno, che per la capacità di includere molteplici metadati interni nel file, è obbligatoria la capacità di leggere tale contenitore da parte di qualunque organizzazione, mentre è fortemente consigliato di adottarlo per la produzione di documenti solo video o audiovisivi — laddove non sia preferibile un formato con gradi di interoperabilità ancora maggiori — quale IMF ad esempio (sotto descritto).

IMF	FORMATO DI PACCHETTO
Nome completo	Interoperable Master Format
Estensione/i	[.pkl|cpl], .xml, .mxf, …
Tipologie MIME	application/xml, /mxf, …
Sviluppato da	Society of Motion Picture and Television Engineers
Tipologia di standard	aperto, estendibile, retrocompatibile, *de iure*
Livello metadati	4
Derivato da	SMPTE Digital Cinema Package
Revisione	2019
Riferimenti	Famiglia di standard **ST2067** della SMPTE: • base: ST2067-1, ST2067-2:2013, ST2067-3:2016, ST2067-5:2012, ST2067-101:2017, ST2067-102:2014, ST2067-200:2018; • audio: ST2067-6:2012, ST2067-8:2013; • applicazioni: ST2067-10, ST2067-20:2013, ST2067-21:2016, ST2067-30:2013, ST2067-40:2017, ST2067-50:2018 • SMPTE RDD47:2019, *Isochronous Stream of XML Documents (ISXD) plugin* • SMPTE Report, *TTML features for IMF Data essence*, 2012 • www.smpte.org • www.imfug.com
Conservazione	Sì, con Applicazioni IMF che usano codec conservabili
Racc. per la lettura	Speciale; obbligatorio per il trattamento e l'interscambio di master e semilavorati cinetelevisivi
Racc. per la scrittura	Speciale; obbligatorio per interscambio, archiviazione e conservazione di contenuti cinetelevisivi.

15. Il formato master interoperabile (IMF) è descritto dalla famiglia di standard **ST2067** della SMPTE ed è nato dall'esigenza condivisa dai maggiori studi cinetelevisivi del mondo di gestire e conservare contenuti audiovisivi in maniera organizzata e attenta all'impatto che tale tipo di documenti, –per effetto delle loro grandi dimensioni informatiche– hanno nei processi di produzione, trasferimento, archiviazione, localizzazione e distribuzione internazionale. Una metodologia che è stata seguita nell'individuazione di tale formato è quella di massimizzare il riutilizzo di porzioni di contenuti audiovisivi che sono identici fra l'oro, archiviate nello stesso formato di pacchetto. Ciò avviene immancabilmente nelle riedizioni del medesimo contenuto (in caso di doppiaggio e altri adattamenti), nonché quando lo stesso master viene trasferito tra più enti durante le fasi di post-produzione, distribuzione e archiviazione.

16. Un documento informatico in tale formato è chiamato 'pacchetto master interoperabile' (*IMP*) e comprende sempre dei contenuti multimediali rappresentati nel tempo lungo una timeline (cfr. §2.10 articolo 1), chiamata *composizione*.
17. IMF utilizza ampiamente le caratteristiche di estendibilità offerte dal formato XML (usato per file sidecar di un IMP, che ne contengono i metadati globali), così come l'interoperabilità della busta MXF e dei suoi schemi operativi, per contenere le essenze audio, video, oltre ai dialoghi e ad altri metadati "temporali" (cioè localizzati cronologicamente in punti specifici della timeline).
18. Un IMP contiene sempre almeno i seguenti file XML:

- un 'elenco di impacchettamento' (PKL, *packing list* in inglese), che contiene il nome completo di un dato pacchetto, l'UID ad esso associato e altri suoi metadati, oltre nome, dimensione, impronta crittografica (cfr. §2.16) e UID associati a ciascuno dei file facenti parte dell'IMP;
- una 'scaletta della composizione' (CPL, *composition playlist* in inglese) che descrive come la timeline del contenuto sia composta a partire dalle risorse (file) indicizzate nella PKL del medesimo pacchetto e, opzionalmente, anche da PKL di altri pacchetti.

19. Un pacchetto master interoperabile è un pacchetto di file in base alla definizione data in §1.1.1, ma la sua CPL, potendo riferirsi anche ad altre PKL oltre alla propria, può descrivere il contenuto di una composizione utilizzando anche contenitori multimediali appartenenti ad altri IMP. Un pacchetto IMF che faccia uso di altre PKL oltre la propria viene chiamato 'pacchetto complementare' (*supplemental package*, in inglese),[52] in quanto dipende da uno o più IMP differenti dal proprio. Il caso d'uso degli IMP complementari è quello di riedizioni o adattamenti successivi del medesimo contenuto, quali:

- doppiaggio (i.e. cambio della lingua parlata);
- aggiunta, o cambio, di *timed text* (i.e. sottotitoli di vario tipo, cfr. § 2.11);
- sostituzione di cartelli e altre parti di video ove vada inserito un testo in una lingua differente (*texted*), o ne vada aggiunto ove non era presente (*textless*);
- riedizione di parti del contenuto audiovisivo per altre finalità di adattamento;
- aggiunta o sostituzione di marche, loghi, o interi titoli di testa o di coda;
- sostituzione dell'intero contenuto con altre codifiche con diverse specifiche tecniche (ad esempio video stereoscopico, HDR/WCG, HFR;[53] audio con diverso numero di canali, ecc.).

20. IMF ammette anche altri tipi di file sidecar opzionali, sempre in XML, quali:

[52] In tal caso la CPL di un pacchetto complementare si dice essere un *versioning file* (*VF*) — mentre invece un pacchetto IMF che si riferisce soltanto alla propria PKL (e qundi ai suoi file MXF), si dice essere –almeno tecnicamente– nella 'versione originale' (*OV*).
[53] Acronimi commerciali che stanno rispettivamente per *high dynamic range*, *wide colour gamut* e *high frame rate*.

- un 'elenco dei profili d'uscita' (OPL, *output profile list* in inglese), che contiene informazioni tecniche su come generare file multimediali auto-consistenti in altri formati di file, incluse informazioni sul formato video (risoluzione, rapporto d'aspetto, ritagli e ingrandimenti) e audio (numero di canali e lingue impiegate) da utilizzare;[54]
- una 'mappa di composizione accessoria' (SCM, *sidecar composition map*), che può racchiudere commenti, annotazioni, ovvero metadati "globali" associabili all'IMP nella sua interezza — particolarmente utile ai fini dell'archiviazione a lungo termine di qualunque tipi di informazione accessoria riguardo al contenuto audiovisivo stesso.

21. Gli eventuali sottotitoli, così come i metadati associabili a particolari punti della timeline (cosiddetti "locali") sono entrambi rappresentati come tracce in un IMP, venendo codificati in XML imbustati ciascuno in file MXF facenti parte del pacchetto. Il dialetto XML dei sottotitoli è sempre TTML o sue ulteriori specializzazioni (cfr. §2.11).
22. Gli altri tipi di metadati locali sono codificabili in file XML senza ulteriori vincoli di specializzazione, associati a particolari istanti della timeline e raggruppati in uno o più buste. Ciascun siffatto contenitore di metadati costituisce un *flusso isocrono di documenti XML* (ISXD). Si raccomanda di avere un file MXF separato per ogni tipologia macroscopica di metadati.[55]
23. Allo scopo di incrementare l'interoperabilità, accanto agli standard che descrivono i componenti generici di IMF, ve ne sono altri che descrivono i suoi schemi operativi, chiamati "Applicazioni". In questi documenti vengono descritti ulteriori vincoli tecnici che armonizzano gli IMP formati per determinate finalità, come ad esempio i codec da usare per le essenze, la loro risoluzione, le specifiche sugli spazi-colore e altre informazioni colorimetriche, la suddivisione delle tracce audio multicanale in "campi sonori", la compressione audio-video, il formato dei sottotitoli, e così via. Molte di queste Applicazioni sono diventate standard di consegna per diversi attori del mercato cinetelevisivo, soprattutto nel campo SVOD (distribuzione a consumo via etere/satellite/internet).
24. Sebbene IMF resti un formato specifico per i settori cinetelevisivi, all'interno di questi è fortemente raccomandata la produzione di documenti in questo formato: sia nei casi di documenti "semilavorati", per i quali siano cioè previsti ulteriori

[54] Bisogna ricordare che, sebbene IMF sia un formato adatto per la post-produzione e l'archiviazione, il suo uso principale è come master cinetelevisivo, da intendersi successivamente convertito in formati specifici agli scopi di post-produzione, distribuzione e conservazione.
[55] Ad esempio 4 file MXF per contenere 4 differenti ISXD: uno con l'elenco dei nomi di attori e/o personalità presenti in ad ogni cambio di scena; uno con informazioni di produzione ad ogni cambio scena (illuminazione, ottiche, parametri delle configurazioni della cinepresa); uno con parametri relativi al restauro fotogramma per fotogramma (difetti fisici della pellicola sia riparati che solo individuati, finestra del fotogramma completo da cui è estratto il fotogramma cinetelevisivo, ecc.); uno con informazioni di postproduzione e visualizzazione in tempo reale (correzione del colore in HDR, parametri degli effetti speciali, coordinate giroscopiche della camera, ecc.).

adattamenti quali montaggio, doppiaggio o altro; sia qualora il contenuto sia destinato ad archiviazione a lungo termine e conservazione. Le PP.AA. di settore sia adoperano per adottare il formato IMF per tali scopi nel minor tempo possibile.
25. Per le organizzazioni che manipolano un elevato numero di IMP, si raccomanda l'uso di un applicativo di gestione dei contenuti che –sfruttando i file sidecar di cui all'articolo 18– consenta una migliore razionalizzazione degli storage contenenti tali contenuti, tramite il tracciamento completo dei file costituenti ogni pacchetto IMP.

MATROSKA	FORMATO CONTENITORE
Nome completo	Matroska
Estensione/i	.mkv, .mka, .mks, .mk3d
Magic number	0x1A45DFA3
Tipo MIME	video/x-matroska, audio/x-matroska
Sviluppato da	comunità open source
Tipologia di standard	aperto (licenza CC BY 4.0), *de facto*, binario
Livello metadati	3
Derivato da	Multimedia Container Format (MCF), Extensible Binary Meta Language (EBML)
Revisione	1.4.9 (20 aprile 2017)
Riferimenti	• www.matroska.org • github.com/Matroska-Org • matroska-org.github.io
Conservazione	No
Racc. per la lettura	Generale con obbligo in lettura
Racc. per la scrittura	Generale; raccomandato per la produzione non cinetelevisiva di contenuti cronologicamente continui

26. Il contenitore Matroska è un altro formato estremamente versatile capace di contenere molteplici essenze di vario tipo. Se ne raccomanda inoltre l'uso in tutti i campi della produzione audiovisiva, salvo nei casi d'uso specifici della post-produzione o sua archiviazione, dove sono preferibili formati contenitori più professionali, quali MXF ovvero MP4.

WEBM	FORMATO CONTENITORE
Nome completo	WebM
Estensione/i	.webm, .weba
Magic number	0x1A45DFA3
Tipo MIME	video/webm, audio/webm
Sviluppato da	Google; On2 Technologies
Tipologia di standard	proprietario (licenza tipo BSD), *de facto*, binario
Livello metadati	1
Derivato da	Matroska, VP9 (video), Vorbis (audio)

Revisione	2018
Riferimenti	• www.webmproject.org • www.matroska.org • www.vorbis.com
Conservazione	No
Racc. per la lettura	Generico; raccomandata la lettura
Racc. per la scrittura	Generico; raccomandata la produzione per brevi clip finalizzate alla distribuzione via internet

27. Il contenitore WebM è ad oggi uno degli standard de facto per la condivisione di contenuti audiovisivi sottoforma di flussi in *streaming* via internet. È in pratica costituito dalla combinazione di una busta Matroska, del codec VP9 per le essenze video e di quello Vorbis per le essenze audio.

MPEG2-TS	FORMATO CONTENITORE
Nome completo	MPEG-2 Transport Stream
Estensione/i	.ts, .m2ts
Magic number	G
Tipo MIME	video/MP2T
Sviluppato da	Moving Picture Experts Group
Tipologia di standard	aperto, estendibile, *de iure*, robusto, binario
Livello metadati	1
Derivato da	MPEG-2
Revisione	2018
Riferimenti	• ISO/IEC 13818-1:2018 • ITU-T Recommendation H.222.0 (2018) • ETSI TS-101-154 v2.6.1 (2019)
Conservazione	No
Racc. per la lettura	Generale raccomandato
Racc. per la scrittura	Speciale; raccomandato per la produzione finale di contenuti cronologicamente continui

MPEG2-PS	FORMATO CONTENITORE
Nome completo	MPEG-2 Program Stream
Estensione/i	.mpg, .mpeg, .vob, .m2p
Magic number	0x0001BA
Tipo MIME	video/MP2P
Sviluppato da	Moving Picture Experts Group
Tipologia di standard	aperto, estendibile, *de iure*, binario
Livello metadati	1
Derivato da	MPEG-2
Revisione	2018

Riferimenti	• ISO/IEC 11172-3:1993 • ISO/IEC 13818-1:2018 • ITU-T Recommendation H.222.0 (2018)
Conservazione	No
Racc. per la lettura	Generale; raccomandato
Racc. per la scrittura	Speciale; sconsigliato per archiviazione/conservazione

28. L'*MPEG-2 Transport Stream* (*TS*) e *Program Stream* (*PS*) sono due formati busta utilizzati per contenere essenze audiovisive, in due scenari di utilizzo diversi:

- MPEG2-TS, introdotto originariamente per una riproduzione sequenziale sotto forma di flusso digitale (cfr. §1.1.1) trasmesso attraverso un canale di comunicazione non affidabile, quale ad esempio il "digitale terrestre";[56] per questo motivo l'MPEG2-TS è dotato di messaggi di verifica dell'integrità legati a meccanismi di protezione del flusso digitale. Adotta inoltre una disposizione interlacciata delle essenze per migliorare le prestazioni dovute allo scenario d'uso della riproduzione sequenziale.
- MPEG2-PS, introdotto originariamente per la distribuzione televisiva aerea o via cavo, è disegnato per contenere flussi audiovisivi digitali (cfr. §1.1.1, con estensioni preferite .mpg e .mpeg), insieme ai metadati necessari al mantenimento della sincronia, al controllo di qualità della banda impiegata e all'organizzazione del canale di comunicazione di tipo televisivo (e.g. supporta il numero e il nome del programma/canale, ecc.). L'MPEG2-PS è disegnato per la riproduzione casuale ed è adottato come contenitore per le evidenze informatiche nei dischi di formato DVD e simili (adottando, in questo caso, l'estensione .vob).

29. Inizialmente questi formati di busta supportavano codec della famiglia MPEG-2, anche se è possibile utilizzare altri tipi di codec.

AVI	FORMATO DI FILE
Nome completo	Advanced Video Interleave
Estensione/i	.avi
Magic number	RIFF
Tipo MIME	video/avi, video/msvideo
Sviluppato da	Microsoft Corporation
Tipologia di standard	proprietario (libero), *de iure*, binario
Livello metadati	3
Derivato da	Resource Interchange File Format (RIFF)
Revisione	2008

[56] Cioè i canali di comunicazione che seguono gli standard di comunicazione denominati **DVB** della ITU e della EBU.

Riferimenti	• Microsoft, *AVI RIFF File Reference*, vs.85 (2008) • RFC-2361 • Hackaday, *AVI File Format*
Conservazione	No
Racc. per la lettura	Generale con obbligo in lettura
Racc. per la scrittura	Generale; sconsigliata la produzione

30. Il formato di busta AVI è una declinazione del più generico formato contenitore RIFF (da cui derivano anche WAVE e AIFF nel campo audio, cfr. §2.9). È stato, storicamente, uno dei primi formati di contenitore destinato a documenti audiovisivi e dunque a supportare l'uso di codec molteplici per le sue essenze. La mancanza di particolari metadati professionali, tuttavia, oltre ad una poca descrizione dei codec impiegati, lo rendono poco adatto ad usi professionali, perciò se ne sconsiglia l'uso per la creazione di nuovi file laddove vi sia un'alternativa possibile.

OGG	FORMATO CONTENITORE
Nome completo	Ogg encapsulated format
Estensione/i	.ogg, .oga, .ogv
Magic number	OggS
Tipo MIME	audio/ogg, video/ogg, application/ogg
Sviluppato da	comunità open source
Tipologia di standard	aperto (licenza tipo BSD), *de iure*, binario
Livello metadati	3
Derivato da	—
Revisione	0 (2003)
Riferimenti	• RFC-3533 • RFC-5334 • xiph.org/vorbis
Conservazione	No
Racc. per la lettura	Speciale; raccomandata lettura
Racc. per la scrittura	Speciale; sconsigliata post-produzione cinetelevisiva

31. Il formato OGG è un contenitore completamente aperto ed estremamente versatile, raccomandato per l'archiviazione di essenze audio, ma al di fuori del caso d'uso cinetelevisivo, dove si raccomandano altri forati di file.

QUICKTIME	FORMATO CONTENITORE
Nome completo	QuickTime™ File Format
Estensione/i	.mov, .qt
Magic number	0x00000000 moov

Tipo MIME	video/quicktime
Sviluppato da	Apple
Tipologia di standard	proprietario, estendibile, *de facto*, deprecato
Livello metadati	4
Derivato da	–
Revisione	2016
Riferimenti	• Apple, *QuickTime File Format Specifications* (2016)
Conservazione	No
Racc. per la lettura	Generale; raccomandata la lettura
Racc. per la scrittura	Generale; sconsigliata la produzione

32. Il contenitore QuickTime, di proprietà della Apple, è stato a lungo uno dei contenitori più usati per le essenze multimediali, grazie soprattutto alla sua estendibilità. Nonostante si tratti di un formato aperto (la Apple pubblica e revisiona continuamente le sue specifiche tecniche), ne detiene la proprietà intellettuale e ne ha solo parzialmente liberalizzato la produzione al di fuori di applicativi o dispositivi hardware di proprietà della Apple o vincolati da accordi di licenza. Per questo motivo, nonostante ne sia fortemente consigliata la capacità di riproduzione da parte di tutte le organizzazioni che utilizzano, elaborano, o producono contenuti audiovisivi, la produzione di nuovi contenuti in questo formato è fortemente sconsigliata (a meno che non sia l'unico formato supportato da specifici applicativi). Inoltre la smpte e la iso hanno derivato dal QuickTime il formato MP4, completamente aperto.

DCP		FORMATO DI PACCHETTO
Nome completo	SMPTE Digital Cinema Package	
Estensione/i	[.pkl|cpl].xml, .mxf, …	
Tipologie MIME	application/xml, /mxf, …	
Sviluppato da	Society of Motion Picture and Television Engineers	
Tipologia di standard	proprietario (libero), estendibile, *de iure*	
Livello metadati	3	
Derivato da	DCP della DCI (Digital Cinema Initiative) e *"InterOp"*	
Revisione	1.3 (2018)	

Riferimenti	Famiglie di standard ('ST'), prassi raccomandate ('RP'), regole tecniche ('EG') **429–433** della SMPTE: • pacchetto: ST429-2:2009, ST429-3:2007, ST429-6:2006, ST429-7:2006, ST429-8:2007, ST429-9:2007 • video: ST429-4:2006, ST428-11:2009, ST428-21:2011, ST428-2:2006, EG432-1:2010; • audio: ST428-1:2006, ST428-2:2006, ST428-3:2006 • sottotitoli: ST429-5:2009, ST429-12:2008, ST428-21:2011; • www.smpte.org • www.dcimovies.com
Conservazione	No (cfr. §2.8 per i sottotitoli)
Racc. per la lettura	Speciale; raccomandato per la distribuzione e post-produzione cinematografica
Racc. per la scrittura	Speciale; raccomandato per distribuzione e (non criptato) per archiviazione di contenuti cinematografici

33. Il formato di pacchetto per il cinema digitale (DCP, *digital cinema package* in inglese) è un pacchetto di file il cui scopo è contenere una copia di contenuto cinematografico da distribuire nelle sale di proiezione. Esistono due tipi di standard DCP: quello della *Digital Cinema Initiatives* (DCI), ora obsoleto, e quello della SMPTE, più recente, che attualmente supporta tutte le novità tecniche.
34. Un DCP è costituito da file multimediali MXF, con schema operativo **OP**2b, contenenti separatamente almeno una traccia audio (in formato PCM non compresso) e una video (con compressione inter-frame di tipo *wavelet*),[57] più alcuni sidecar file in XML. Da questo punto di vista, l'architettura del pacchetto è simile a quella del formato di master interoperabile (IMF, sopra descritto). Ciò che rende le DCP peculiari sono i vincoli ai formati audiovisivi delle sale teatrali; la possibilità, mediante messaggi inseriti in specifici istanti della timeline della composizione, di comandare operazioni in sala quali il controllo del proiettore, del sistema di tende, sipari e luci, o di altri dispositivi di intrattenimento proprietari; la possibilità di cifrare l'intero contenuto sbloccandolo soltanto con una chiave crittografica (cfr. §2.16) generata in maniera sicura, tipicamente dallo stesso applicativo che ha masterizzato la DCP.
35. Le DCP cifrate hanno infatti tutti i loro file MXF cifrati, ciascuno mediante crittografia simmetrica (algoritmo AES) con *chiave-contenuto* a 256 bit. La chiave-contenuto di ciascun MXF è sempre archiviata e distribuita cifrata da una chiave pubblica (RSA), la cui corrispondente chiave privata è confidenzialmente mantenuta

[57] La compressione impiegata, visivamente priva di perdite (*visually lossless* in inglese), è la medesima utilizzata dall'algoritmo JPEG2000, cfr. §2.7.

da dispositivi certificati dalla DCI — tipicamente dispositivi di masterizzazione o riproduzione del contenuto (attaccati ai proiettori delle sale cinematografiche), ai quali viene distribuita l'evidenza informatica contenente la chiave-contenuto da sbloccare per riprodurre il DCP. Tale evidenza è formata in un particolare file chiamato KDM (cfr. §2.16).

36. L'uso della crittografia per proteggere i DCP pone tuttavia diverse problematiche in merito all'archiviazione a lungo termine e alla conservazione dei documenti cinetelevisivi in tale formato: in seguito a smarrimento o a scadenza naturale dei certificati a chiave pubblica, infatti, non sarebbe più possibile generare altre KDM che potrebbero sbloccare la DCP.

37. Mentre da una parte si raccomanda che le organizzazioni dedicate alla produzione e riproduzione di opere cinematografiche compiute siano in grado di riprodurre almeno le DCP *non* cifrate, si sconsiglia alle medesime organizzazioni di produrre DCP, salvo per le seguenti finalità:

- DCP *non* criptate per finalità di archiviazione a lungo termine e conservazione del master dedicato alla proiezione cinematografica;
- DCP *criptate* per la sola finalità di distribuzione nelle sale cinematografiche.

85. Si sconsiglia di produrre DCP secondo lo standard DCI (ovvero secondo lo standard *de facto* e intermedio fra i due qui elencati, denominato "InterOp"), favorendo invece lo standard *de iure* SMPTE in quanto è l'unico a supportare, mediante aggiornamenti lenti ma costanti delle specifiche, le migliorie tecnologiche delle sale cinematografiche, sia in campo audio, che video, che intrattenimento di sala in generale.

39. Si raccomanda alle organizzazioni che archiviano un contenuto in formato DCP di archiviarne anche una versione in un formato indipendente dalle caratteristiche tecniche di una sala cinematografica digitale, quali quelli individuati in questo stesso capitolo.

DCDM		FORMATO DI PACCHETTO
Nome completo	Digital Cinema Distribution Master	
Estensione/i	.tif/.tiff, .wav	
Tipologie MIME	image/tiff, sound/wav	
Sviluppato da	Society of Motion Picture and Television Engineers	
Tipologia di standard	aperto, estendibile, retrocompatibile, *de iure*	
Livello metadati	2	
Derivato da	Tagged Information File Format	
Revisione	1.3 (27 giugno 2018)	

Riferimenti	Famiglia **ST428** di standard/prassi SMPTE: • video: ST428-1:2006, ST428-5:2010, ST428-9:2008, RP428-6:2009; Adobe® TIFF™ Revision 6.0 (1992) • audio: ST428-2:2006, ST428-3:2006, ST428-4:2010, ST428-3:2006; EBU – Tech 3285-1 2.0 (2001) • sottotitoli: ST428-7:2010, ST428-10:2008, ST429-5: 2009; ST428-5:2010 • www.smpte.org
Conservazione	Sì (cfr. §2.8 per i sottotitoli)
Racc. per la lettura	Speciale; obbligatorio per la post-produzione e distribuzione cinematografica
Racc. per la scrittura	Speciale; raccomandato per l'archiviazione di contenuti per la distribuzione cinematografica, previa riversamento in pacchetto IMF Application #4

40. Il master per la distribuzione del cinema digitale (DCDM) è un altro formato basato su sequenze di fotogrammi, ove l'essenza audio è costituita da file WAVE monaurali (uno per ciascun canale stereo, 5.1 ovvero 7.1) mentre l'essenza audio è costituita da file TIFF adeguatamente formati (cfr. §2.6), che codificano l'immagine nello spazio-colore DCI *X'Y'Z'*, basato su *code-value* a numeri interi e sulla colorimetria riferita alla luce proiettata su uno schermo bianco, che però, da un punto di vista teorico, comprende ogni colore visibile da un "osservatore standard". Il pacchetto DCDM, così come il DCP sopra descritto, esistono in due versioni: uno primo standard *de facto* della DCI e il più recente standard *de iure* della SMPTE, cui ci si raccomanda di attenersi.
41. Sebbene la lettura del DCDM sia obbligatoria per le organizzazioni che si occupano della post-produzione e della masterizzazione per il cinema digitale, nonché fortemente consigliata per l'archiviazione a lungo termine dei contenuti cinematografici, si consiglia altresì la produzione dei futuri master cinematografici nel formato IMF, che presenta caratteristiche di maggior interoperabilità con le filiere contemporanee di distribuzione cinematografica, oltre a supportare un maggior numero di metadati e di resilienza alle riedizioni.
42. Si consiglia il riversamento di tutti i pacchetti DCDM in pacchetti IMF Application #4 (cosiddetta "*Cinema mezzanine*").

CINEMADNG		FORMATO DI FILE
Nome completo	Adobe® CinemaDNG™	
Estensione/i	.dng; .wav	
Specializzazione di	TIFF/EP	
Tipo MIME	video/x-adobe-dng	
Sviluppato da	Adobe Systems	

Tipologia di standard	proprietario (brevettato), estendibile, *de iure*, binario
Livello metadati	2
Derivato da	Adobe® DNG
Revisione	1.1.0.0 (2011)
Riferimenti	• Adobe, *CinemaDNG image data format specification* v1.1.0.0, 2011 • Adobe, *Digital Negative (DNG) Specification* v1.4.0.0, 2012
Conservazione	No
Racc. per la lettura	Speciale; fortemente raccomandato per archiviazione, conservazione e post-produzione fotocinematografica
Racc. per la scrittura	Speciale; fortemente consigliato per archiviazione/ conservazione fotografica; consigliato riversamento su IMF in caso cinematografico

43. Il formato CinemaDNG™ è una specifica aperta della Adobe per rappresentare il video professionale sotto forma di sequenze di fotogrammi, ove ciascun file è rappresentato come DNG.

D.I. BASATO SU EXR		FORMATO DI PACCHETTO
Nome completo	Digital Intermediate	
Estensione/i	.exr; .wav	
Tipologie MIME	image/exr, sound/wav	
Sviluppato da	major film studio	
Tipologia di standard	aperto, retrocompatibile, *de facto*	
Livello metadati	2	
Derivato da	–	
Revisione	–	
Riferimenti	• www.smpte.org	
Conservazione	Sì	
Racc. per la lettura	Speciale; obbligatorio in post-produzione cinetelevisiva	
Racc. per la scrittura	Speciale; fortemente raccomandato per l'archiviazione cinetelevisiva; consigliato riversamento in IMF	

D.I. BASATO SU DPX		FORMATO DI PACCHETTO
Nome completo	Digital Intermediate	
Estensione/i	.dpx; .wav	
Tipologie MIME	image/x-dpx, sound/wav	
Sviluppato da	Technicolor; Kodak	
Tipologia di standard	aperto, retrocompatibile, *de facto*, deprecato	
Livello metadati	2	

Derivato da	–
Revisione	–
Riferimenti	–
Conservazione	No
Racc. per la lettura	Speciale; obbligatorio in post-produzione cinetelevisiva
Racc. per la scrittura	Speciale; consigliato per archiviazione di scansioni da pellicola cinematografica; consigliato riversamento in IMF, ovvero in DI basato su EXR

44. Il nome *Digital Intermediate* (*DI*) indica il processo di post-produzione cinematografica digitale che ha soppiantato l'analogo processo fotochimico su pellicola da 16mm, 35mm o 65mm. Agli albori della post-produzione digitale, quando la fotografia era prevalentemente su pellicola negativa e la proiezione nelle sale basata su pellicole positive, il processo DI cominciava con la scansione digitale del negativo già sviluppato –fotogramma per fotogramma– e terminava con la stampa digitale (*film-out*, in inglese) su pellicola internegativa o interpositiva di pari calibro, da cui venivano poi stampate le copie positive in 35mm o 70mm per la distribuzione. Con l'avvento predominante delle cineprese digitali da un lato del processo e i proiettori per il cinema digitale dall'altro, il termine di è oggi diventato semplice sinonimo di post-produzione interamente digitale. È rimasta comunque l'esigenza di memorizzare i contenuti video in pacchetti di file a "sequenza di fotogrammi" (cfr. §2.12) in quanto ciò rende più semplice la modifica di singole scene, parti di scene e persino fotogrammi individuali — sia dal punto divista dello storage ove il contenuto è memorizzato, che delle prestazioni degli applicativi impiegati. Per gli stessi motivi, le sequenze di fotogrammi non utilizzano compressione, o al più utilizzano algoritmi *lossless* (dunque, considerando la sequenza di fotogrammi come un unico flusso video, si tratta sempre di compressione intra-frame, cfr. §2.10). La suddivisione in "rulli" (cioè clip audiovisive della durata di 15–25 minuti circa[58]) è di solito rispettata dividendo le sequenze di fotogrammi in sottocartelle del medesimo pacchetto

45. L'audio viene invece codificato in uno o più file WAVE (cfr. §2.9), tipicamente monoaurali, uno per canale. Non esistono veri e propri documenti che descrivono tali specifiche, così come la loro archiviazione a lungo termine (storicamente in nastri digitali LTO mediante il formato di archiviazione TAR, cfr. §2.13): si tratta, dunque, di uno vero e proprio standard *de facto*.

46. Per quanto riguarda il formato di file usato per i singoli fotogrammi, il DI tradizionale ha da molto tempo adottato il formato DPX (creato dalla Kodak® come evoluzione del Cineon™ e standardizzato poi dalla SMPTE, cfr. §2.6), in quanto tale formato permette nativamente la rappresentazione digitale della densitometria di

[58] Considerando un framerate di 24fps tipico del cinema, questo equivale dai 20'000 ai 30'000 fotogrammi.

una pellicola negativa scansionata, ovvero di una positiva stampata, oltre a supportare metadati tipici della postproduzione cinematografica.[59] Il formato DPX tuttavia, nonostante sia attualmente sotto revisione da parte di SMPTE, non è più adeguato agli standard di archiviazione multimediale moderni, a causa della poca efficacia nel mantenere metadati inerenti alle annotazioni o alla codifica colorimetrica. Per questo e altri motivi, come ad esempio le esigenze tecniche della computer-grafica (*CG*) e degli effetti speciali digitali (*VFX*), i pacchetti DI contemporanei hanno abbandonato il formato DPX in favore di OpenEXR (anch'esso introdotto in §2.6).

47. Si raccomanda a tutti gli enti operanti nel settore della post-produzione cinematografica di poter leggere pacchetti di file di entrambe i tipo, anche se si sconsiglia fortemente la produzione di nuove sequenze di fotogrammi in formato DPX (in favore del formato OpenEXR). Per la produzione di nuovi pacchetti destinati all'archiviazione o conservazione, si raccomanda inoltre l'impiego dell'IMF (sopra descritto), per la sua maggiore capacità di interoperabilità e riusabilità.

AMF	FORMATO DI PACCHETTO
Nome completo	ACES Metadata File
Estensione/i	.amf
Tipologie MIME	application/amf+xml
Sviluppato da	Academy of Motion Picture Arts and Sciences
Tipologia di standard	aperto, estendibile, retrocompatibile, testuale
Livello metadati	3
Derivato da	ACESclip
Revisione	1.2 (2019)
Riferimenti	Standard e bollettini tecnici dell'AMPAS: • TB-2014-009, *ACES clip-level metadata file format* • S-2014-006, *A common file format for Look-Up Tables* • TB-2014-012, *ACES version 1.0 component names* • TB-2014-010, *design, integration and use of ACES LMT* • www.oscars.org/aces • www.acescentral.com
Conservazione	Sì
Racc. per la lettura	Speciale; consigliato in post-produzione cinetelevisiva

[59] Quali il TimeCode (cfr. §2.12) e il KeyCode™: quest'ultimo, in pratica, un UID che si associa al singolo fotogramma –addirittura della singola perforazione– al numero di serie di ogni rullo, insieme alla sua marca ed emulsione fotochimica.

Racc. per la scrittura	Speciale; consigliato in post-produzione, archiviazione e conservazione di contenuti cinetelevisivi in ACES

48 Il formato ACESclip è un dialetto XML riservato ad un file *sidecar* che possa accompagnare diversi tipi di file video, all'interno di un processo di trasporto, elaborazione o archiviazione di contenuti che rispetti, colorimetricamente, il sistema ACES introdotto in §2.6. In particolare si consiglia di affiancare un file ACESclip (estensione .ACESclip.xml) sia a file con contenuti compatibili alle specifiche colorimetriche ACES (in pratica, il solo spazio-colore ACES2065-1 descritto nello standard ST2065-1).

XDCAM		FORMATO DI PACCHETTO
Nome completo	XDCAM™ package	
Estensione/i	.mxf, .xml, …	
Tipologie MIME	application/mxf, /xml, …	
Sviluppato da	Sony Corporation	
Tipologia di standard	proprietario (chiuso), *de facto*	
Livello metadati	4	
Derivato da	–	
Revisione	–	
Riferimenti	–	
Conservazione	No	
Racc. per la lettura	Speciale; raccomandato in post-produzione e distribuzione di contenuti cinetelevisivi	
Racc. per la scrittura	Speciale; sconsigliata la produzione	

49. Il formato di pacchetto proprietario XDCAM è utilizzato da alcune cinecamere della Sony Corporation, anche se il nome XDCAM™ si riferisce a una gamma ben più ampia di prodotti per il video professionale, inclusi dispositivi di archiviazione (*Professional Disc*, ovvero dischi P2) e di riproduzione. Come formato di pacchetto, XDCAM prevede una ramificazione in sottocartelle predefinite, l'uso di vari codec video (e.g. DV, MPEG-2 parte 2, MPEG-4 ovvero M4V), più alcuni file XML di contorno, che ne descrivono i metadati e il pacchetto d'insieme. A causa delle licenze d'uso legate al formato, se ne sconsiglia fortemente l'uso per la produzione di nuovi contenuti (salvo laddove il capitolato tecnico di consegna lo preveda come unica opzione per una data qualità oggettiva). Tuttavia, la sua ampia diffusione come standard *de facto* impone di considerarlo come formato da poter aprire da parte delle organizzazioni che si occupano di post-produzione o masterizzazione dei contenuti.

2.12.1 Raccomandazioni per la produzione di documenti

1. Per quanto riguarda la produzione di documenti audiovisivi in generale, come verrà anche ribadito nel capitolo sul riversamento (§3.3), si raccomanda una valutazione di interoperabilità su tutti i formati di tutte impiegati in un documento informatico multimediale. In particolar modo, qualora il contenuto audiovisivo sia rappresentato da un pacchetto di file (ove, tipicamente, ciascun file descrive metadati, singole essenze o loro porzioni) andrà valutata l'interoperabilità del formato del pacchetto, seguito dall'interoperabilità di ogni file potenzialmente contenuto in esso. Per ciascun file andrà valutata l'interoperabilità del formato contenitore (wrapper, cfr. §1.1.1), se del caso, la tipologia di essenze in esso rappresentate e dunque l'interoperabilità dei codec usati da ciascuna essenza nel contenitore.
2. Relativamente ai soli formati di contenitori e pacchetti multimediali oggetto di questa sezione, si fanno le seguenti raccomandazioni:

- Per contenuti audiovisivi, prevalentemente di natura cinetelevisiva o comunque con destinazioni d'uso professionali –siano essi master per la distribuzione tramite i canali della televisione tradizionale o OTT– da sottoporre o meno a procedimenti di archiviazione o conservazione, si raccomanda particolarmente il formato master interoperabile (IMF). Tale formato di pacchetto di file consente, tecnicamente, una separazione delle essenze (anche relativamente a edizioni multiple e localizzazione); operativamente, può trasportare anche la separazione del contenuto in sue parti funzionali (e.g. marche e loghi iniziali, prologo, titoli di testa, "capitoli", titoli di coda, ecc.). In entrambe i casi, tali separazioni possono essere sfruttate in caso di riedizione, trasporto e archiviazione, allo scopo di minimizzare tempi, risorse computazionali e occupazione di storage, riducendo i costi infrastrutturali e migliorando al contempo le caratteristiche archivistiche del contenuto. Si raccomanda, tuttavia, di formare i pacchetti IMF secondo una specifica "Applicazione" (anch'esse standardizzate dalla SMPTE) relativa al dato contenuto, in modo da estendere l'interoperabilità con le aziende di settore che le adottano. Il formato IMF può invece non essere ideale quando si verificano una o più delle seguenti condizioni:
    - sia necessario conservare le essenze audiovisive senza perdita di qualità della compressione,
    - non via sia un'Applicazione IMF adeguata al processo, o
    - sia necessaria una modifica del contenuto per cui è preferibile l'utilizzo di sequenze di fotogrammi (cfr. qui sotto).

- Per contenuti sempre con destinazioni d'uso professionale (prevalentemente cinetelevisivi) ove sia necessario preservare la massima qualità video (tramite assenza di compressione o l'uso di compressione *lossless*) e, al contempo, è preferibile un approccio basato su sequenze di fotogrammi, si raccomanda l'uso di DI basato su EXR.
- Per particolari casi d'uso cinetelevisivi, vincolati a specifiche regole tecniche, è possibile adottare altri formati, come ad esempio:
    - il pacchetto di cinema digitale (DCP) per la distribuzione finale dei film nelle sale cinematografiche;[60]
    - il DI basato su sequenze di DPX per contenuti scansionati da pellicole cinematografiche e non ulteriormente elaborati;
    - l'MPEG2 per la sola trasmissione dei contenuti come flussi in tempo-reale (e.g. televisione via cavo, satellitare, ovvero OTT).
- Per tutti gli altri tipi di contenuto, si raccomanda l'utilizzo di un solo file contenitore che racchiuda tutte le essenze e i metadati necessari comunque alla riproduzione del documento audiovisivo. In particolar modo, si raccomanda l'utilizzo del *wrapper* MXF, in quanto potenzialmente integrabile con funzionalità avanzate che rendono facilmente riutilizzabile un contenuto generico anche in ambiti professionali.
- Per contenuti audiovisivi generici, ove sia prioritario facilitarne la riproduzione da parte di applicativi di larga diffusione (anziché dotarlo di un contenitore professionale che supporti metadati interoperabili e una migliore facilità di riversamento), si raccomanda in particolar modo il *wrapper* MP4.

2.13 Archivi compressi

1. ***Nota Bene***: Alcuni dei pacchetti di file oggetto del presente Allegato sono di fatto creati, come ultimo passaggio, impacchettando una porzione di filesystem (cfr. §1.1.2) 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.
2. 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).

[60] Si sconsiglia di utilizzare il DCP come formato per archiviazione e conservazione per via della compressione jpeg2000 delle essenze video e, soprattutto, a causa della crittografia a chiave mista pubblica/privata con cui vengono formate le copie per le sale e che rischia, senza una corretta gestione delle chiavi crittografiche (cfr. §2.16), di rendere inutilizzabile il documento archiviato in DCP.

3. Un tipico esempio ove sono impiegate entrambe le convenzioni è rappresentato quando il documento, originariamente costituito da più file, è riversato in un unico file mediante due passaggi distinti e consecutivi: prima viene "pacchettizzato" in un unico file; successivamente tale file viene compresso usando un formato a scelta tra una pluralità di formati (e i corrispondenti algoritmi di compressione). 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.

TAR	FORMATO CONTENITORE
Nome completo	UNIX Standard Tape Archive (TAR)
Estensione/i	.tar
Magic number	ustar
Tipo MIME	application/x-tar
Sviluppato da	comunità open source
Tipologia di standard	aperto, retrocompatibile, *de iure*, binario
Livello metadati	1
Derivato da	PKZIP
Revisione	7 (2017)
Riferimenti	• IEEE 1003.1:2017, POSIX *base specifications, issue 7* • GNU, Basic Tar Format, (2017)
Conservazione	Sì, ma dipende dal contenuto della busta TAR
Racc. per la lettura	Generico; obbligatoria la lettura
Racc. per la scrittura	Generico; raccomandata la produzione per archivi su nastri digitali generici o file applicativi Linux/UNIX

ZIP	FORMATO CONTENITORE
Nome completo	Zip
Estensione/i	.zip, .zipx
Magic number	PK 0x[0304|0506|0708]
Tipo MIME	application/zip
Sviluppato da	PKWARE®
Tipologia di standard	proprietario (libero), retrocompatibile, *de iure*, binario
Livello metadati	3
Derivato da	PKZIP

Revisione	6.3.5 (2018)
Riferimenti	• ISO/IEC 21320-1:2015 • PKWARE, *.ZIP File Format Specification*, v6.3.5 (2018)
Conservazione	Sì, ma dipende dal contenuto della busta ZIP
Racc. per la lettura	Generico; obbligatoria dalla versione 6.3.1 (2007)
Racc. per la scrittura	Generico; raccomandata la versione 6.3.1 o precedenti

GZIP	FORMATO CONTENITORE
Nome completo	GNU Zip
Estensione/i	.gzip
Magic number	0x01F8B
Tipo MIME	application/gzip
Sviluppato da	comunità open source
Tipologia di standard	aperto (GNU LGPL), retrocompatibile, *de iure*, binario
Livello metadati	1
Derivato da	PKZIP
Revisione	6.3.5 (2018)
Riferimenti	• RFC-1952 • RFC-6713 • www.zlib.org/rfc-gzip.html
Conservazione	Sì, ma ipende dal contenuto della busta GZIP
Racc. per la lettura	Generico; obbligatorio
Racc. per la scrittura	Generico; raccomandato per archivi contenenti applicativi Linux/UNIX

7-ZIP	FORMATO CONTENITORE
Nome completo	*7-Zip* compressed archive format
Estensione/i	.7z
Magic number	7z 0xBCAF271C
Tipo MIME	application/x-7z-compressed
Sviluppato da	Igor Pavlov
Tipologia di standard	aperto (GNU LGPL), retrocompatibile, *de facto*, binario
Livello metadati	3
Derivato da	compressione Lempel-Ziv-Markov (LZMA)
Revisione	18.06 (2018)
Riferimenti	• Pavlov I., LZMA SDK (2018)
Conservazione	Sì, ma ipende dal contenuto della busta 7-Zip

Racc. per la lettura	Generico; consigliato
Racc. per la scrittura	Generico; consigliato

RAR	FORMATO CONTENITORE
Nome completo	Roshal Archive file format
Estensione/i	.rar; .r[00–99]
Magic number	Rar! 0x1A0700, Rar! 0x1A070100
Tipo MIME	application/java-archive
Sviluppato da	Eugene (algoritmo) e Alexander (applicativi) Roshal
Tipologia di standard	aperto (lettura), proprietario (scrittura), retrocompatibile, *de facto*
Livello metadati	3
Derivato da	compressione Lempel-Ziv-Storer-Szymański (LZSS)
Revisione	5.61 (2018)
Riferimenti	• www.rarlab.com • theunarchiver.com
Conservazione	No
Racc. per la lettura	Generico; altamente sconsigliato; valutare opportunità di riversamento in altro formati di archiviazione
Racc. per la scrittura	Generico; Sconsigliato.

JAR	FORMATO CONTENITORE
Nome completo	Java Archive file format
Estensione/i	.jar
Magic number	0x5F27A889
Tipo MIME	application/jar-archive
Sviluppato da	Verizon Media; Oracle Corporation
Tipologia di standard	proprietario (libero), retrocompatibile, *de facto*, binario
Livello metadati	3
Derivato da	ZIP
Revisione	2018
Riferimenti	• Oracle, JAR File Overview, 2018 • Oracle, JAR File Specification, 2018
Conservazione	Sì, ma dipende dal contenuto della busta JAR
Racc. per la lettura	Specifico; raccomandato in ambito ICT e sviluppo codice
Racc. per la scrittura	Specifico; raccomandato in ambito ICT e sviluppo codice in linguaggio Java

ISO	FORMATO CONTENITORE
Nome completo	Immagine di volume ISO9660
Estensione/i	.iso
Magic number	CD001 (dal 32768,mo 34817,mo o 36865mo byte)
Tipo MIME	application/x-iso9660-image
Sviluppato da	International Organization for Standardization
Tipologia di standard	proprietario (libero), retrocompatibile, *de iure*, binario
Livello metadati	1
Derivato da	
Revisione	6.3.5 (2018)
Riferimenti	• ECMA-119, 3rd Ed. (2017) • ISO/IEC 13490-1:1995, […]: *General* • ISO/IEC 13490-2:1995, […]: *Volume & file structure* • ECMA-168, 2nd ed., 1994 • Famiglia di standard **13346** della ISO/IEC • ECMA-167, *Universal Disk Format (UDF)*, 3rd ed., 1997
Conservazione	Sì, ma dipende dal contenuto della busta ISO9660
Racc. per la lettura	Specifico; fortemente raccomandato in ambito ICT
Racc. per la scrittura	Specifico; raccomandato in ambito ICT per immagini normali e forensi di dispositivi di *storage* a blocchi

4. Questo standard descrive prevalentemente due oggetti che possono a livello logico contenersi l'una nell'altra:

1. un contenitore generico per la memorizzazione "byte per byte" di un'evidenza informatica costituente il volume di un dispositivo di *storage* a blocchi (tipicamente ad accesso casuale);
2. un filesystem (cfr. §1.1.2), la cui specifica è codificata nello standard deprecato ISO 9660;

4. Lo standard ISO 9660, obsoleto, è stato sostituito, sia come storage che come filesystem virtuale, da altri due standard: l'ISO 13490, che è prevalentemente un suo aggiornamento; l'ISO 13346, denominato *Universal Disk Format* (UDF), che introduce invece contenitore e filesystem nuovi. In passato, diverse estensioni furono introdotte –come standard *de facto* però– da alcune organizzazioni: le Apple ISO9660 Extensions (per il filesystem proprietario HFS+), l'IEEE P1282 ("Rock Ridge"), il Microsoft® "Joilet", l'IBM "El Torito". Come filesystem invece, l'ISO9660 (che continua a dare il nome al formato di file) e le sue estensioni furono originariamente impiegati per i dischi ottici di tipo CD, DVD e BluRay (BD) — nelle versioni "ROM"

a sola-lettura o meno. Il formato, rappresentato in un unico file è stato utilizzato per distribuire immagini trasportabili di tali dischi, allo scopo di masterizzarli in un secondo momento. In una terza fase temporale, si è consolidato come standard *de facto* per conservare immagini "pseudoforensi" di versioni installabili o preinstallate di alcuni software, simulandone la distribuzione nell'obsoleto supporto informatico basato su disco.

VMDK	FORMATO CONTENITORE E DI PACCHETTO
Nome completo	Virtual Machine Disk Format
Estensione/i	.vmdk
Magic number	KDMV 0x[01|02]000000
Tipo MIME	application/x-vmdk
Sviluppato da	VMware (Dell Incorporated)
Tipologia di standard	aperto, retrocompatibile, *de facto*, binario o testuale
Livello metadati	3
Derivato da	ISO 660
Revisione	VMFS-5 (2011)
Riferimenti	• VMware, *Virtual Disk Format* (2011) • VMware, Virtual Disk Development Kit (VDDK) v6.7.1
Conservazione	Sì, ma dipende dal contenuto della busta VMDK
Racc. per la lettura	Usare SOS o applicativi commerciali nei reparti IT.
Racc. per la scrittura	Usare SOS o applicativi commerciali nei reparti IT.

5. Il formato VMDK, proprietario VMware ma standard *de facto*, permette di rappresentare uno storage a blocchi in formato "sparso" (cioè rappresentando solo i blocchi effettivamente inizializzati dal dispositivo), in uso da parte di macchine virtuali e container. È un formato sia di contenitore che di pacchetto perché è disponibile in due varianti: nella prima è un singolo file che contiene sia l'header con i metadati interni, sia il contenuto a blocchi vero e proprio dello storage virtualizzato. Nella seconda variante, invece, è costituito da un file di testo contenente i soli metadati più uno o più file binari contenenti solo la rappresentazione blocco per blocco di parti del dispositivo.

DMG	FORMATO CONTENITORE
Nome completo	Apple® Disk Image
Estensione/i	.dmg
Specializzazione di	Immagine di volume ISO9660
Tipo MIME	application/x-apple-diskimage
Sviluppato da	Apple Incorporated

Tipologia di standard	proprietario (libero), retrocompatibile, *de facto*, binario
Livello metadati	3
Derivato da	Apple New Disk Image Format (NDIF, .img)
Revisione	–
Riferimenti	• newosxbook.com/DMG.html • Apple, *File System Programming Guide* (2018)
Conservazione	No
Racc. per la lettura	Nessuna raccomandazione
Racc. per la scrittura	Valutare riversamento in altri formati per dati semplici. Conservare/archiviare i file .dmg in caso di backup di applicativi nativi Apple (macOS®, iOS®, ecc.).

2.13.1 Raccomandazioni per la produzione di documenti

1. 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 descritti 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.

2.14 Documenti amministrativi

1. Sono qui elencati alcuni formati di file utilizzati per documenti amministrativi di utilizzo generale, da parte delle PP.AA. e di altri enti, su tutto il territorio nazionale, quali:

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

2. Si precisa che, per lo svolgimento di procedimenti amministrativi non contemplati in questa sezione, le PP.AA. posso utilizzare anche altri formati di file descritti sia nel resto di questo Allegato, sia non compresi da esso 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]

FATTURAPA		FORMATO DI FILE
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 • *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	

3. 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,

[61] Si veda in tal senso, per quanto riguarda il protocollo informatico le Linee guida cui questo Allegato afferisce, per quanto riguarda gli schemi di identificazione elettronica e i servizi fiduciari elettronici il §2.16 del presente Allegato, che rimanda alla normativa attualmente in vigore.

- descrizione dei beni o servizi oggetto della fatturazione da un punto di vista fiscale,
- metadati inerenti all'eventuale trasporto fisico dei beni oggetto della fatturazione.

4. È anche prevista la possibilità di allegare alla fattura uno o più allegati che vengono codificati [62] e "immersi" all'interno del file (che diviene così un contenitore).

CDA2		FORMATO DI FILE
Nome completo	Clinical Document Architecture	
Estensione/i	.xml	
Specializzazione di	XML	
Tipo MIME	application/xml	
Sviluppato da	HL7 International	
Tipologia di standard	proprietario (libero), estendibile, *de iure*, testuale	
Livello metadati	2	
Derivato da	HL7 CDA Rel. 1.0	
Revisione	2.0 (2018)	
Riferimenti	• ISO/HL7 27932:2005, *HL7 Clinical Document Architecture* • www.fascicolosanitario.gov.it/Standard-documentali • Foglio di stile per la consultazione interoperabile • Circolare AGID N°4/2017	
Conservazione	Sì	
Racc. per la lettura	Specifico; consultare la normativa in materia	
Racc. per la scrittura	Specifico; consultare la normativa in materia	

5. Il CDA è uno standard di markup basato su XML pensato per lo scambio informatico di documenti clinici ed è, inoltre, uno degli standard di riferimento per il Fascicolo Sanitario Elettronico (FSE), la cui normativa si può trovare sul sito ufficiale del FSE: www.fascicolosanitario.gov.it. Il CDA è uno standard definito originariamente da ANSI e successivamente sviluppato dall'organizzazione senza scopi di lucro Health Level 7. Il CDA specifica la sintassi e fornisce una struttura di base (data set di riferimento) per realizzare l'intera semantica di un documento clinico (guida implementativa). Un documento CDA è in grado di contenere qualunque tipo di informazione clinica; supporta inoltre testo non strutturato e può incorporare documenti nei formati PDF, DocumentML e RichText Format, così

[62] Per la precisione, la codifica binaria è ottenuta mediante algoritmo *Base64,* specificato in RFC-4648.

come immagini di tipo JPEG e PNG. La versione 2 del CDA è stata adottata dallo standard ISO/HL7 27932 del 2005.

SEGNATURA DI PROTOCOLLO		FORMATO DI FILE
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 presenti Linee guida	

6. La segnatura di protocollo è descritta nell'Allegato 6 delle presenti Linee guida.

ASSERZIONE SPID		FORMATO DI FILE
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 • 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), siano esse in linguaggio SAML (basato su XML), ovvero in formato JWT (basato su JSON), cfr. §2.3, 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.

2.15 Applicazioni e codice sorgente

1. Esistono una miriade di formati per codificare le applicazioni (altrimenti dette, impropriamente anche se a causa di una forzatura storica, *binari*). Tipicamente, tali formati dipendono dall'architettura dell'hardware e dal SO impiegato. Per tale motivo figurano tanto gli eseguibili Microsoft® (.exe, .com, .msi), quanto quelli macOS® (.pkg), quanto le *app* per dispositivi mobili, sia Android (.apk), che iOS® (.ipa). Oltre agli eseguibili, le porzioni di codice statico (.a, .lib) e dinamico (.so, .dll, .dylib) sono anch'esse dotate di formati specifici, largamente dipendenti dal SO.
2. La distinzione si semplifica nel caso dei codici sorgente scritti nei vari linguaggi di programmazione –interpretati o compilati– per i quali la tipologia di file è in realtà unica: un semplice file di testo (con codifica dei caratteri ASCII ovvero in qualche variante di UNICODE o UTF), la cui estensione (e, di conseguenza, anche il tipo MIME) da un'indicazione logica circa il linguaggio di programmazione o di *scripting* in cui sono codificate le proposizioni al suo interno.
3. Né il formato dei file contenenti le applicazioni, né i linguaggi di programmazione sono oggetto delle Linee guida di cui il presente Allegato è parte integrante. Si rimandano le PP.AA. alle *Linee guida sul riuso del software* per gli obblighi e le raccomandazioni in materia di sviluppo applicativo.

2.16 Applicazioni crittografiche

1. Per meglio chiarire il contesto, si premette che i servizi fiduciari elettronici sono normati da fonti di livello pari o superiore a quello delle presenti Linee guida, le quali contengono obblighi e raccomandazioni anche in merito ai formati di file utilizzati per tali servizi. Tali fonti, includenti le loro successive modificazioni, sono:

- Regolamento (UE) Nº 910/2014 del Parlamento europeo e del Consiglio (regolamento "eIDAS");
- Decisioni di Esecuzione (UE) collegate con il regolamento eIDAS;
- articoli 20, 21, 24, 28, 35 e 36 del CAD;
- *Linee Guida contenenti le Regole Tecniche e Raccomandazioni afferenti la generazione di certificati elettronici qualificati, firme e sigilli elettronici qualificati e validazioni temporali elettroniche qualificate*, emanate con Determinazione AgID Nº121/2019;
- D.P.C.M. del 21 maggio 2013 (*Nuove Regole Tecniche Firma Elettronica Avanzata*).

2. I documenti informatici normati dalle suddette fonti, utilizzati per servizi fiduciari elettronici e altri servizi che impiegano la crittografia, al fine di aumentare le caratteristiche di confidenzialità, integrità, autenticità e non ripudio di evidenze informatiche, si classificano, generalmente, in:

a) buste crittografiche contenenti firme, sigilli o validazioni temporali elettroniche ed, eventualmente, i documenti informatici stessi a cui tali servizi fiduciari afferiscono;
b) certificati elettronici di creazione di firma, sigillo o validazione temporale elettronica;
c) certificati elettronici di autenticazione di siti web (WAC);
d) certificati elettronici di attribuzione o autenticazione;
e) certificati elettronici di certificazione;
f) richieste di sottoscrizione di certificato elettronico;
g) liste di revoca o di fiducia (di certificati elettronici);
h) buste contenenti chiavi crittografiche;
i) buste contenenti documenti cifrati elettronicamente;
j) buste contenenti impronte crittografiche;

3. Nel caso in cui la busta crittografica contenga solo la firma elettronica (o le firme elettroniche), senza il documento stesso a cui essa o esse si riferiscano logicamente, si parla di firma *detached*; quando invece la busta crittografica contiene sia il documento che la sua firma elettronica si parla di firma *enveloping*; quando, infine, il documento informatico costituisca esso stesso busta crittografica per la firma apposta, si parla di firma *enveloped*. La stessa distinzione nominativa si applica, *mutatis mutandis*, al sigillo elettronico ovvero alla validazione temporale elettronica.
4. Viene aggiunta una specifica relativa allo standard *de facto* dei file contenenti una o più impronte crittografiche in un formato puramente testuale (ASCII) e alcuni formati relativamente al trasporto di informazioni crittografiche in ambiti specifici.

CHECKSUM		FORMATO DI FILE
Nome completo	impronta crittografica	
Estensione/i	.sha2, .sha1, .md5, .ripemd160, …	

Magic Number	–
Sviluppato da	–
Tipologia di standard	aperto, *de facto*, testuale
Livello metadati	1
Derivato da	XAdES
Revisione	–
Riferimenti	• SANS, *An introduction to file integrity checking on UNIX systems*, GIAC paper, 2003
Conservazione	No
Racc. per la lettura	Generale; obbligo di lettura come normale file di testo
Racc. per la scrittura	Generale; raccomandato la produzione di un'impronta *detached* così costituita in assenza di altri meccanismi di verifica dell'integrità impliciti nei formati o non collegati con essi da vincoli logici più forti e robusti.

5. Per quanto concerne le impronte crittografiche (anche dette *digest* o *thumbprint* in inglese) e le chiavi crittografiche, ove esse non siano utilizzate per finalità, ovvero non siano contenute in file il cui formato non è altrimenti definito da altre fonti normative, si raccomanda di adottare la seguente metodica per creare e archiviare un digest *detached* di file (il cui nome generico sia *nomefile.ext*):

a) scegliere una funzione di *hash* crittografico adeguatamente robusta (si raccomanda SHA-256 o superiore, come definito in RFC-6234 ovvero da Avvisi pubblicati da AgID) e rappresentare il nome dell'algoritmo mediante una stringa di caratteri alfanumerici minuscoli *hash*: senza interruzioni, caratteri di spaziatura o altri simboli (ad esempio, sha256 per SHA-256, ripemd160 per RIPEMD a 160 bit, md5 per MD5 e così via);
b) calcolare l'impronta crittografica dell'intero file con la funzione di *hash* di cui al punto 1;
c) creare un file chiamato *nomefile.ext.hash* (che avrà dunque estensione *.hash*), possibilmente nella medesima cartella ove si trova il file stesso;
d) salvare in questo file, mediante codifica ASCII a 7-bit (cfr. RFC-2045) la sola rappresentazione esadecimale (con cifre minuscole) dell'impronta di cui al punto 2; non andranno aggiunti altri caratteri (incluse spaziature e a-capo — *newline* in inglese), né prima né dopo l'impronta;
e) per quanto possibile, mantenere la località di referenza tra il file originale e quello contenente la sua impronta crittografica mediante i metodi esposti nel §1.1.2, cioè mantenendo sempre file e impronta nella medesima cartella, con la medesima *naming convention* definita ai punti 1 e 2 di questo elenco numerato.

KDM		FORMATO DI FILE
Nome completo	Key Delivery Message	
Estensione/i	.kdm.xml	
Specializzazione di	XML	
Tipo MIME	application/kdm+xml	
Sviluppato da	Society of Motion Picture and Television Engineers	
Tipologia di standard	aperto, retrocompatibile, *de iure*, testuale	
Livello metadati	4	
Derivato da	XML Digital Signature	
Revisione	—	
Riferimenti	Famiglia di standard **ST430** della SMPTE: • ST430-1:2006, *D-Cinema operations – KDM* • ST430-2:2006, *D-Cinema operations – Digital Certificate* • W3C Recommendation *XML Signature Syntax and Processing*, 2002 (deprecato)	
Conservazione	No	
Racc. per la lettura	Specifico; raccomandato per le sale cinematografiche	
Racc. per la scrittura	Specifico; raccomandato nella produzione dei master cinematografici per la distribuzione nelle sale	

6. Il formato DCP è un formato di busta contenente le cosiddette "chiavi di contenuto" nel contesto del cinema digitale (per i cui formati audiovisivi vedasi §2.12), cioè chiavi simmetriche AES a 256 bit di lunghezza), fino ad uno specifico server Digital Cinema, cui è concessa la facoltà di riprodurre il contenuto di un determinato pacchetto di cinema digitale (DCP) entro un intervallo di tempo finito e predeterminato. Per ogni DCP cifrato, infatti, le essenze di ciascun file MXF sono cifrate ciascuna con una chiave di contenuto. La sicurezza del DCP si ottiene cifrando ulteriormente tutte le chiavi di contenuto di un dato DCP mediante crittografia asimmetrica (RSA a 2048 bit di lunghezza), ove ciascun server è dotato di un HSM contenente sia la chiave privata della coppia (e il relativo certificato di creazione di sigillo elettronico avanzato) che un orologio indipendente anti-manomissione per la validazione temporale elettronica. Su tutti i KDM si appone un sigillo *enveloping* che protegge da contraffazione la specifica DCP da riprodurre (mediante sia l'UID del pacchetto stesso che le impronte crittografiche dei singoli file che lo costituiscono, cfr. §1.1.1), l'intervallo temporale durante il quale ne è concessa la riproduzione e, soprattutto, la copia cifrata delle sue chiavi di contenuto.

3 Raccomandazioni sui formati di file

1. I formati da utilizzare nell'ambito delle presenti Linee guida sono quelli previsti nel §2 del presente Allegato. Nello scegliere tali formati di file, da utilizzare per i propri documenti informatici, le organizzazioni possono effettuare la valutazione di interoperabilità (§3.1) che prediliga formati aperti, non proprietari, standard *de iure*, estendibili, parlanti, completamente robusti, indipendenti dal dispositivo –secondo la classificazione fatta in §1.2.2– al fine di:

- prevenire il rischio della loro obsolescenza tecnologica (rischio che incrementa con il divergere delle caratteristiche del formato da quelle sopraelencate);
- mitigare il rischio del *vendor lock-in*,[63] che sussiste soprattutto per i formati proprietari, non codificati in standard *de iure*, o per quelli dipendenti dal dispositivo;
- facilitare il più possibile un loro futuro riversamento in altro formato, prediligendo caratteristiche quali l'estendibilità, la codifica testuale, la compatibilità in avanti e, in misura minore, la robustezza.

2. I formati di file, di contenitori, di pacchetti di file e di codec elencati nel §2 rappresentano, in generale, una famiglia allargata, con candidati più o meno adatti ad un particolare scopo. Nel caso dei codec video, allo scopo di indicare esempi di formati non interoperabili da utilizzare il meno possibile per conservazione o archiviazione, sono elencati anche alcuni codec chiusi o con particolari limitazioni d'uso dovute a licenze, brevetti o royalties, allo scopo di indicare esplicitamente le organizzazioni che avessero documenti multimediali con tali codec verso un riversamento o altra procedura di gestione documentale mirata.
3. Così come accade per i sistemi di archiviazione digitale (storage) –e per tutta la tecnologia in generale– non si può pensare che un formato di file, seppur revisionato nel tempo, possa essere perennemente attuale né di uso corrente. L'obsolescenza dei formati di file e dei dispositivi di archiviazione può essere mitigata soltanto mediante *riversamenti periodici* da una tecnologia verso un'altra; questo vale tanto per le tecnologie di storage che per i formati dei file.
4. Per questo motivo, si consiglia che le PP.AA. e le imprese che hanno necessità di archiviazione e riuso a lungo e lunghissimo termine dei documenti informatici, pianifichino in anticipo una strategia di selezione degli *storage* e dei formati di file che sia il più possibile congiunta, in modo da minimizzare il numero di riversamenti relativi a ciascuno di questi due fattori (ottimizzando la continuità operativa e, nel complesso, riducendo i costi derivanti da tali operazioni).

[63] Consiste, in questo caso, nell'impossibilità tecnologica e/o giuridica di elaborare file (inclusa la conversione in altri formati) senza l'ausilio del proprietario del formato: avvalendosi di sue consulenze, beni, tecnologie, proprietà intellettuale, licenze software o altro.

3.1 Valutazione di interoperabilità

1. La valutazione di interoperabilità in merito ai formati dei file, prevista dalle Linee guida di cui il presente Allegato è parte integrante, può essere redatta da qualunque ente pubblico o privato tratti documenti informatici. La valutazione di interoperabilità è redatta con cadenza annuale dalle PP.AA. che trattino documenti informatici in formati diversi da quelli di cui al presente Allegato, ovvero conformi a questi formati ma utilizzati disapplicando gli obblighi e raccomandazioni ivi contenuti.
2. Allo scopo di effettuare la valutazione di interoperabilità, le organizzazioni effettuano una ricognizione di tutte le loro procedure amministrative e/o processi di business, allo scopo di individuare ogni tipologia di documenti informatici trattati (o trattabili). Il valore di questa ricognizione può andare ben al di là degli scopi di cui alle presenti Linee guida.
3. La valutazione di interoperabilità consiste in un dettagliato rapporto circa le seguenti azioni (obbligatorie se indicate in **grassetto**):

a) Includere nell'attività di classificazione (cfr. §1.2.2) un **censimento dei formati** di file e delle tipologie di storage attualmente utilizzati, con particolare riferimento a quelli non elencati nel §2 del presente Allegato.
b) Per ciascun formato di file adottato, elencare tutti i **dettagli tecnici** dei medesimi, quali ad esempio:
    - nome dei formati e, laddove applicabile, dei dialetti, profili, codec, schemi operativi;
    - suddivisione tra formati generici e specifici (cfr. §1.2.3);
    - versioni utilizzate nei documenti già esistenti, ovvero producibili dagli attuali applicativi;
    - altre caratteristiche tecniche non vincolate dalle specifiche di cui ai punti precedenti (e.g. lingue adottate nei testi, numero di canali audio, spazi-colore, risoluzione per immagini e video, *bitrate* massimo, algoritmi di cifratura, presenza di password, ecc.).
c) Elencare i **processi di riversamento** di formato attualmente in corso nell'organizzazione, con particolare riferimento ai software applicativi impiegati e alle procedure tecniche (automatiche, semiautomatiche o completamente manuali) adottate per configurare tali riversamenti, con lo scopo prioritario di rendere tali riversamenti riproducibili.
d) Elencare le motivazioni attuali che hanno portato alla scelta di ciascun formato di file per il trattamento dei documenti informatici. In particolar modo, se del caso, distinguere i formati di file tra quelli adottati per i documenti:

- accettati "in entrata" dal pubblico ovvero da altre organizzazioni,
- utilizzati ad uso esclusivamente interno,
- pubblicati, ovvero prodotti "in uscita" verso altre organizzazioni,
- archiviati ovvero mandati in conservazione.

e) Valutare l'esistenza di standard o di iniziative di standardizzazione a livello internazionale, europeo e nazionale, relativamente alle tipologie di documenti informatici trattati.
f) Quantificare l'eventuale **necessità** di operare sui medesimi documenti informatici nell'arco di una finestra temporale futura.
g) Valutare gli scenari ove successive modificazioni o revisioni dei documenti vengano prodotte in formati diversi da quello originale.
h) Valutare la sussistenza di leggi o altri tipi di obblighi in merito alla **conservazione delle evidenze informatiche nel formato originale** di acquisizione o formazione.
i) Valutare i formati di file di **categorie specifiche**, nonché l'opportunità di riutilizzo dei documenti informatici di ciascuna classe (come da punto 1) da parte di PP.AA. e organizzazioni che operano al di fuori dello specifico settore per il quale il formato e i suoi scenari d'utilizzo sono stati prefigurati.
j) Dipendenza dei formati di file da:
    - licenze d'uso, marche e brevetti o altra **proprietà intellettuale**,
    - sistemi e architetture **proprietarie**, o comunque,
    - sistemi e architetture che, pur senza i suddetti vincoli, sono comunque associati a costi di manutenzione ordinaria o straordinaria, senza la quale diviene a rischio o è fortemente ridotta la capacità di elaborare i suddetti documenti.
k) Inserimento dell'obsolescenza dei formati di file e delle tecnologie di archiviazione all'interno di una più ampia strategia di trasformazione digitale dell'organizzazione.

Si faccia riferimento al Glossario delle presenti linee guida per la definizione dei termini non ulteriormente introdotti in questo Allegato (e che sono qui indicati in colore azzurro Italia).

4. La valutazione di interoperabilità, in quanto parte della gestione documentale, andrebbe effettuata periodicamente allo scopo di individuare tempestivamente cambiamenti delle condizioni espresse dai punti sopra elencati e permettere quindi di valutare eventuali azioni. È proprio nella ripetizione della valutazione di interoperabilità ad intervalli regolari che divengono manifesti i formati in via di obsolescenza (quando non già conservati in un formato obsoleto).

3.2 Indice di interoperabilità

1. Allo scopo di coadiuvare la valutazione dei formati di file relativamente all'interoperabilità, l'Agenzia propone innanzi tutto un modello semplificato e quantitativo ove, ad ogni formato di file (anche quelli non elencati nel §2 di questo Allegato), viene associato un valore numerico. Tale valore assegnato al formato è dato dalla somma dei valori associati a tutte le caratteristiche di cui ai punti a)–g) del capitolo §1.2.2.
2. Nel caso di formati di pacchetti o contenitori, andrà fatta una valutazione per ogni componente e considerato come addendo il valore più basso (cioè peggiore) per ciascuna delle sue componenti.
Per un formato contenitore andrà valutato sia il formato della busta che il formato del suo contenuto, ad esempio per i contenitori multimediali (cfr. §2.12), il formato dei codec di ciascuna essenza ivi contenuta; per un formato di pacchetto andranno valutati i formati di tutti i file compresi nel pacchetto e, qualora il pacchetto comprenda dei file contenitori, valutarli con il medesimo criterio sopra esposto.
3. Sia per i formati contenitori che per quelli di pacchetto verrà usato come addendo, per caratteristica, il valore numerico peggiore tra quelli dei suoi file (o dei codec di ciascun file). Si prenda, a titolo di esempio, l'indice di interoperabilità di uno specifico pacchetto di file ove tutti i file utilizzino formati indipendenti dal dispositivo tranne un file multimediale, la cui busta e tutte le sue essenze sono anch'esse indipendenti dal dispositivo tranne una sola essenza audio (che ad esempio potrebbe richiedere la presenza di un dispositivo hardware proprietario per la sua decodifica). Il pacchetto andrà considerato comunque come "dipendente dal dispositivo" e quindi gli verrà applicato un modificatore pari a 0 (anziché a +3), proprio a causa della presenza dell'unico file il cui formato è dipendente dal dispositivo.
4. Considerando una scala che va dal file più interoperabile (con indice pari a 20) a quello meno interoperabile (indice pari a 0), un valore pari a 12 o superiore può essere considerato sufficiente dal punto di vista dell'impatto di tale formato relativamente ad interoperabilità e obsolescenza; valori inferiori indicano problematiche oggettive che vanno affrontate il prima possibile in ottica di riversamento o altre metodologie.
5. Si precisa che l'indice di interoperabilità è meramente indicativo, in quanto non tiene conto di alcuna peculiarità che un'organizzazione possa avere in merito a specifici formati di file. Ad esempio, l'utilizzo di formati chiusi, proprietari o dipendenti dal dispositivo dovrebbe essere sempre un fattore di esclusione almeno nella scelta di formati da usarsi per la formazione di nuovi documenti informatici.

3.3 Riversamento

1. Il riversamento di formato, precedentemente introdotto, comporta il trasferimento di un documento informatico in un formato di file diverso: contestualmente, ciò può comportare anche una duplicazione da un sistema di storage ad un altro (volume, disco, nastro, filesystem, storage ad oggetti o altri). Quando si parla di riversamento di file, s'intende un riversamento di almeno il suo formato. Per quanto già detto nell'introduzione al §3, le strategie relative alla gestione dei formati e dei sistemi di storage vanno spesso di pari passo; perciò, può essere in alcuni casi operativamente ed economicamente vantaggioso effettuare i riversamenti di formato contestualmente a quelli di storage. I riversamenti vengono pianificati a seguito di una nuova valutazione di interoperabilità, considerando i costi e i benefici di una tale operazione.
2. Quando si effettua un riversamento finalizzato alla conservazione del file si può e, in certi casi previsti dalla legge, si deve, conservare anche la copia del file nel formato originario. In entrambe questi casi il file originario è conservato indipendentemente dal suo formato di file originario, purché sia conservata anche –in una forma logicamente e univocamente legata ad esso– copia conforme del medesimo in un formato adatto alla conservazione (tra quelli individuati al §2 del presente Allegato). Le considerazioni in questo e nei successivi paragrafi si applicano non solo ai formati di file, ma anche, *mutatis mutandis*, ai formati di buste, pacchetti di file, flussi binari e codec. Lo scopo del riversamento congiunto alla conservazione del documento nel formato originario è consentire la convalida di firme, sigilli o validazioni temporali elettroniche che, eventualmente già presenti nel documento originario, non potrebbero essere riportate –senza invalidarle– nel documento riversato.
3. Quando si effettua un riversamento finalizzato alla conservazione, il file riversato è una copia digitale di un documento digitale e, come tale, la conformità della copia è attestata in base alla normativa vigente, inclusa la certificazione di processo come riportata nell'Allegato 3 delle presenti Linee guida.
4. È importante che nella scelta dei nuovi formati di file (e, *mutatis mutandis*, di contenitori, pacchetti di file, flussi digitali e codec), così come nella scelta metodologica circa l'esecuzione del riversamento, si considerino le peculiarità tecniche del formato sorgente e di quello riversato, con particolare riferimento sia alla perdita di dati e metadati, sia alla diversa qualità o rappresentazione tecnica dei medesimi.
5. Quando si effettua un riversamento massivo finalizzato alla conservazione di più documenti informatici –a prescindere se sia conservato o meno il documento nel suo formato originario– il processo di riversamento include i passaggi indicati nel successivo elenco numerato. Tali passaggi –particolarmente 1 e 2– possono essere

inquadrati come parte di un processo certificato di conformità della copia riversata (cfr. Allegato 3 delle presenti linee guida).

1. Il riversamento di un formato viene effettuato mediante un processo certificato che ne garantisce l'integrità (effettiva o per lo meno sostanziale, come intesa nel punto 2) e la riproducibilità. Ogni procedura di riversamento è descritta nel manuale di gestione e, se del caso (cfr. §3.1), nella valutazione di interoperabilità, in tutti i suoi dettagli tecnici, inclusi quelli sui formati di destinazione e quelli riversati.
2. Per le PP.AA., il riversamento avviene sempre in formati che ne migliorano l'interoperabilità, o comunque non la peggiorano (come, stabilito, ad esempio, mediante il calcolo dell'indice di interoperabilità), tenendo conto degli obblighi sulla specificità di formato introdotti in §2.2. In particolar modo *non* sono opportuni, relativamente alla classificazione di formati di cui al §1.2.2, i seguenti riversamenti, da un formato:
    - aperto verso formati chiusi,
    - non proprietario, ovvero proprietario a libero utilizzo, verso formati proprietari,
    - non dipendente dal dispositivo verso formati dipendenti da dispositivo,
    - formato parlante verso formati muti.
3. Per ogni file riversato di un processo massivo automatico ovvero semiautomatico il processo produce un'attestazione del riversamento specifico di quel file, ove le attestazioni circa documenti riversati come parte della medesima procedura vengono collezionati in un registro di riversamento, che contiene (globalmente rispetto alla particolare procedura e individualmente per ogni file elaborato) almeno:
    a. un riferimento temporale opponibile a terzi ai sensi dell'art. 41 del DPCM 22 febbraio 2013 relativo all'inizio o alla fine del riversamento (indicando chiaramente a quale dei due tempi ciascun riferimento temporale si riferisca);
    b. indicazioni sul sistema informativo impiegato (per esempio: nome, numero di revisione sia del sistema operativo che del software; nome della macchina e suoi indirizzi di rete o altri numeri identificativi unici delle componenti hardware; identificativi unici del software quali i numeri di licenza; nome o identificativo unico e orario di accesso al sistema operativo dell'utenza sotto cui l'applicativo ha agito);
    c. nome del file sorgente, posizione nel filesystem e metadati esterni (cfr. §1.1.2);

d. formato sorgente del file, sua versione del formato e metadati interni (cfr. §1.1.3);
e. impronta crittografica del file sorgente;
f. nome del file riversato, posizione nel filesystem e metadati esterni;
g. formato di riversamento, sua versione e metadati interni convertiti dal punto d;
h. impronta crittografica del file destinazione;
i. in caso di file contenitori (come sorgente o destinazione), i metadati ai punti d, g[64] si intendono riferiti alla busta, cui si aggiunge elenco completo del contenuto della medesima (p.es. essenze e i codec impiegati per ciascuna di esse, con i loro eventuali metadati e profili); deve inoltre essere previsto –nei casi ove sia tecnicamente possibile– file con imbustamento nidificato (cfr. §1.1.1);
j. in caso di pacchetti di file (come sorgente o destinazione), i controlli c–h[65] si intendono riferiti a ciascun file componente il pacchetto, cui si aggiunge l'indicazione dei metadati deducibili dall'intero pacchetto nella sua interezza, laddove non esplicitamente descritti negli eventuali file-manifesto (cfr. §1.1.1);
k. eventuali errori tecnici, anomalie o ambiguità riscontrate durante il riversamento.

4. Nel rispetto delle leggi sulla privacy in vigore (D.Lgs. N°101 del 10 agosto 2018, del GDPR e loro successive modificazioni), il riversamento in altro formato di file costituisce un'ulteriore occasione per ottemperare agli obblighi in materia di adeguatezza, pertinenza, minimizzazione ed esattezza dei dati personali ivi contenuti, così come della liceità del loro trattamento e della loro eventuale pseudonimìa.
5. Qualora sussistano obblighi di legge (come nel caso della protezione dei dati personali o della conservazione sostitutiva) o altri tipi di vincoli nel preservare l'evidenza informatica costituita dal documento nella sua interezza, il documento nel formato originale viene conservato *insieme* a un suo riversamento in formato più interoperabile. Tale associazione logica deve anch'essa essere scritta nel registro di riversamento (ad esempio associando le impronte crittografiche dei due come indicate ai punti e,h[66] del punto 2).
6. Salvo in casi in cui siano applicabili le considerazioni di cui ai punti 3 o 4, un riversamento di formato altera l'evidenza informatica, intaccandone dunque l'integrità da un punto di vista strettamente tecnico; esistono tuttavia riversamenti di formato che mantengono il contenuto documentale *sostanzialmente* invariato (inclusi quelli discussi nei punti 3 o 4), costituendo

[64] Leggasi «al punto d più al punto g».
[65] Leggasi «i controlli dal c all'h».
[66] Leggasi «al punto e più al punto h».

una valida possibilità per il riversamento. Tali possibilità di riversamento, qualora siano individuate, sono comprese nel manuale di gestione documentale. In esso viene descritto come il contenuto documentale viene sostanzialmente preservato, incluse le considerazioni sulla similitudine tra le due evidenze informatiche. Sono incluse anche una o più delle seguenti trasformazioni che, se applicabili, sono effettuate durante il riversamento:

- In caso il documento nel formato originale venga mantenuto per assolvere ad obblighi di legge o altro, specificare come i due file (l'originale e il riversato) vengano logicamente associati fra loro (come descritto al punto 4).
- Laddove le evidenze vengano modificate da algoritmi non reversibili (p.es. compressione con perdita) un'analisi puntuale o statistica dell'ammontare di informazione persa durante il riversamento (p.es. misurata in SNR minimi e massimi tra l'essenza sorgente e quella destinazione), con riferimento a standard di misura riconosciuti a livello internazionale, europeo o nazionale.
- In caso di perdita di metadati interni, indicare quali metadati si perdono e come essi vengano comunque riportati nel registro di riversamento.
- In caso metadati convertiti biunivocamente da un formato all'altro (incluso il caso in cui più metadati nel formato sorgente siano oggetto di accorpamenti o separazioni in altri metadati nel formato di destinazione), descrivere gli algoritmi che sono implementati per ciascun metadato.[67]
- In caso di metadati interni convertiti in maniera *non* invertibile, descrivere gli algoritmi impiegati (come al punto precedente), aggiungendo esplicitamente quali informazioni si perdono e come esse siano trasformate irreversibilmente.[68]
- In caso il formato destinazione ammetta dei metadati obbligatori che non hanno un analogo nel formato sorgente (ovvero ammetta metadati facoltativi che si ritiene comunque opportuno inserire nel

[67] A titolo di esempio, sia dato un campo 'Scadenza' nel formato sorgente e un campo 'NotAfter' nel formato destinazione, le cui sintassi sono, rispettivamente, il numero intero di giorni passati da un'*epoca* pari al 1 gennaio 2019 (fuso orario italiano: CEST) e una stringa di testo formattata come da RFC-3339 (in fuso orario UTC). Viene dunque specificato l'algoritmo, perfettamente invertibile, che mappa un valore di "42" in "2019-02-14T01:00:00Z".

[68] A titolo di esempio, sia dato un campo 'dimensione1' nel formato sorgente e un campo 'Width' nel formato destinazione. Nel manuale di riversamento, andrà giustificata la motivazione per cui un valore di "-42.58cm" nel formato sorgente, dotato di segno e unità di misura, viene mappato in "426" indicato implicitamente in millimetri e arrotondato al numero intero più vicino. È anche opportuno che sia indicato il motivo per cui un metadato relativo ad una dimensione (spaziale?) non meglio identificata sia (sempre?) mappato in un campo riservato alla 'larghezza' di qualcosa.

formato riversato, anche se non hanno un analogo nel formato originario), specificare come gli vengono assegnati valori.

7. In alcuni casi,[69] imbustando un documento informatico in un contenitore aggiuntivo potrebbe mantenere integro il contenuto all'interno, perciò un riversamento in tal senso –cioè un semplice imbustamento del documento senza alterarlo– può essere reversibile. L'integrità del file nella sua interezza è tuttavia compromessa. Analogamente reversibile è invece includere un documento rappresentato da un file all'interno di un pacchetto di file. Entrambi le metodologie non risolvono i problemi di interoperabilità e obsolescenza originali per cui il riversamento è stato concepito, ma possono mitigarli, soprattutto se vengono scelti contenitori o pacchetti parlanti che compensano la mancanza di metadati nel formato originale.
8. Analogamente al punto 3, in alcuni casi, un documento che viene re-imbustato (cfr. §1.1.1), nel senso che il riversamento consiste nel sostituire la busta che lo conteneva con un'altra, potrebbe mantenere, in alcuni casi, l'integrità del contenuto della busta. Verrebbe persa comunque l'integrità della busta e, come per il punto 3, l'integrità del file nella sua interezza. Tuttavia questa metodologia potrebbe risolvere un problema di obsolescenza e interoperabilità legato al solo formato della busta usata in precedenza.

6. Quando si effettui un riversamento *non* finalizzato alla conservazione è comunque opportuno considerare i passaggi elencati al precedente punto 5, ad eccezione dei punti 1 e 2 del medesimo elenco, che si riferiscono al riversamento qualora inquadrato in processi di conservazione o certificazione di processo.

[69] Alcuni formati contenitori, imbustando un documento, si limitano ad aggiungervi una busta costituita da una o più evidenze informatiche in punti predeterminati del documento o essenza originale (tipicamente in testa e in coda, ma anche inseriti in parti centrali). Quando avviene tecnicamente un procedimento di questo tipo l'evidenza informatica imbustata è alterata *reversibilmente*, potendo quindi essere ricostruita bit a bit. Alcuni formati di contenitore di file, tuttavia, effettuano alterazioni irreversibili del contenuto, che dunque non rimane integro a seguito di un'operazione di "sbustamento" (in inglese, *unwrapping*).