

Formati file previsti dal Sistema di Conservazione Digitale di Maggioli spa

1 INTRODUZIONE

Il presente documento è l'allegato al Manuale del Servizio di conservazione digitale, dedicato alla descrizione dei formati (di rappresentazione di file) ammessi dalla quarta "generazione" delle *Descrizioni Archivistiche*, a cui le specifiche tecniche di versamento in conservazione di Maggioli spa fanno riferimento.

Rispetto alla versione precedente, questo documento prevede la (potenziale) trasmissione in conservazione di tutti i formati citati nell'allegato 2 delle linee guida AgID di riferimento e di alcuni altri che, nel tempo, hanno trovato ampia diffusione.

NB: una ulteriore, importantissima, differenza sta nel fatto che ora sono ammessi (NON rifiutati o bloccati) anche documenti rappresentati in formati NON idonei alla conservazione a lungo termine:

- in questo documento ne ricordiamo i criteri di selezione, in modo che i Produttori possano comportarsi di conseguenza, in fase di formazione o accettazione dei documenti in ingresso (in gestione corrente; es. a Protocollo), limitando i casi anomali
- Maggioli spa assicura la leggibilità e intellegibilità dei soli documenti rappresentati in formati idonei alla conservazione a lungo termine; per gli altri formati file, quest'onere torna nelle mani del Soggetto Produttore dei documenti trasmessi al sistema di conservazione
- Maggioli spa comunque assicura per tutti i documenti conservati il mantenimento delle caratteristiche di autenticità, integrità, immutabilità e reperibilità previste da AgID

Sommario

1 INTRODUZIONE	1
2 Riferimenti Normativi e Ambito di Applicazione	3
2.1 Altre fonti e Norme di riferimento	3
2.2 Versionamento e aggiornamento delle specifiche tecniche del Servizio	4
3 CRITERI DI SELEZIONE DEI FORMATI	5
3.1 TITOLARITÀ	6
3.2 APERTURA	6
3.3 TRASPARENZA	6
3.4 ROBUSTEZZA	6
3.5 STABILITÀ	6
3.6 AUTO-CONTENIMENTO	7
3.7 AUTO-DOCUMENTAZIONE	7
3.8 INDIPENDENZA DAL DISPOSITIVO	7
4 ELENCO DEI FORMATI AMMESSI DAL SISTEMA	8
4.1 Tabella dei formati	9
5 SPECIFICHE DI RAPPRESENTAZIONE	12
5.1 Mime-Type previsti	12
5.2 Riversamento di formato	15

2 Riferimenti Normativi e Ambito di Applicazione

L'adozione dei formati qui descritti si fonda sul quadro normativo italiano in materia di documenti e archivi informatici. I riferimenti principali sono le Linee Guida sulla formazione, gestione e conservazione dei documenti informatici (Maggioli 2021) dell'Agenzia per l'Italia Digitale (AgID).

Tali Linee Guida, adottate ai sensi dell'Art. 71 del Codice dell'Amministrazione Digitale (CAD), hanno carattere vincolante (*valenza erga omnes*) e definiscono le regole tecniche per la corretta gestione documentale, dalla formazione sino alla conservazione.

In particolare, il presente documento si focalizza sull'applicazione dell'Allegato 2 "Formati" delle Linee Guida, che elenca i formati file che possono maggiormente garantire il principio dell'interoperabilità tra i sistemi di gestione documentale e conservazione, in base alla normativa vigente, riguardante specifiche tipologie di documenti (documenti informatici e documenti amministrativi informatici).

Una corretta gestione dei documenti, sin dalla fase di formazione, è la migliore garanzia per l'adempimento degli obblighi di natura amministrativa, giuridica e archivistica.

2.1 Altre fonti e Norme di riferimento

- Determinazione AgID n. 407/2020 del 9 settembre 2020
- Digital Formats for Library of Congress Collections¹
- ISO - International Organization for Standardization
- PRONOM (PRactical Online cOMpendium of file formats technical registry)²

¹ <https://www.loc.gov/preservation/resources/rfs/format-pref-summary.html>

² <https://www.nationalarchives.gov.uk/PRONOM/Format/proFormatSearch.aspx?status=new>

2.2 Versionamento e aggiornamento delle specifiche tecniche del Servizio

Ogni aggiornamento di questa specifica tecnica determina una variazione delle Descrizioni Archivistiche utilizzate dal Sistema di conservazione.

Variazioni particolarmente impattanti sono notificare via PEC o Report ai titolari degli archivi di conservazione coinvolti; in questo caso si tratta di una notifica di variazione o adeguamento dei flussi di versamento richiesti.

Le nuove specifiche si applicano primariamente ai nuovi flussi di versamento attivati. La notifica, riguardante la necessità di aggiornare i flussi di versamento, è trasmessa solo quando i flussi obsoleti vengono posti in dismissione (c.d. “deprecati”).

Facendo rimando al Manuale (Disciplinare) del Servizio, per tutto quanto non qui riportato, le Descrizioni Archivistiche (DA) sono la rappresentazione delle configurazioni applicate, in modo omogeneo, ad una porzione di Archivio e ai flussi, di gestione e conservazione, a queste indirizzate.

Tramite il Report di dettaglio dell’Archivio di Conservazione, fornito periodicamente (e a richiesta) da Maggioli spa, è possibile ottenere una evidenza storicizzata delle Descrizioni Archivistiche utilizzate nel tempo per ogni tipologia documentaria utilizzata.

3 CRITERI DI SELEZIONE DEI FORMATI

Allo scopo di poter garantire nel tempo la fruibilità e intelligibilità delle informazioni contenute nei documenti oggetto del servizio, rispetto alla totalità dei file utilizzabili in “gestione corrente”, sono stati definiti da AgID opportuni criteri di selezione dei formati conservabili, cui si applicano la valutazione di interoperabilità, accessibilità e il **calcolo dell'indice di interoperabilità**, che definiscono la compatibilità rispetto al processo di conservazione, nonché la necessità di eventuali, successivi, interventi di riversamento.

Per una descrizione dettagliata si rimanda all'allegato 2 delle Linee guida AgID, “3.1 Valutazione di interoperabilità”, di cui riportiamo di seguito i criteri utilizzati.

Tali criteri potranno essere usati anche dai Soggetti Produttori per definire le regole di selezione dei formati da implementare in capo al sistema di gestione documentale versanti, in particolare rispetto a quelli destinati alla conservazione a lungo termine.

Questi criteri riguardano particolari caratteristiche dei formati di rappresentazione (in file) dei documenti informatici/elettronici e, in particolare:

- Proprietà/Titolarità (licenza)
- Apertura
- Trasparenza
- Robustezza
- Stabilità
- Auto-contenimento
- Auto-documentazione
- Indipendenza dal dispositivo

3.1 TITOLARITÀ

Preferenza per formati **non proprietari**.

Un formato si dice *proprietario* quando è legato all'esistenza di un'azienda o di un'organizzazione privata che ne detiene i diritti di proprietà intellettuale e ne gestisce a proprio piacimento le specifiche.

Invece, un formato è definito *non proprietario* quando la gestione delle sue specifiche è condivisa nell'ambito di una comunità di sviluppatori o di un ente di standardizzazione, ed è quindi **libero da restrizioni legali sul suo utilizzo**, nonché dal controllo di interessi privati.

3.2 APERTURA

Preferenza di formati **aperti**: si parla di formato aperto quando le sue specifiche sono pubbliche, liberamente accessibili ed esaustive.

3.3 TRASPARENZA

Tiene conto del grado di semplicità con cui è possibile la **fruizione dei contenuti** codificati secondo uno specifico formato utilizzando semplici strumenti di base.

3.4 ROBUSTEZZA

Un formato è robusto quando **in caso di corruzione del file** (perdita accidentale di uno o più bit/byte) consente il recupero, totale o parziale, dei suoi contenuti

3.5 STABILITÀ

Un formato è più o meno stabile a seconda della **frequenza** con cui subisce nel tempo delle modifiche e variazioni (le quali comportano spesso continue migrazioni e di conseguenza possibili errori o perdite).

3.6 AUTO-CONTENIMENTO

Un formato è auto-contenuto quando consente di includere tutte le risorse necessarie per la sua rappresentazione (es. un documento di testo deve includere i font utilizzati, così come devono essere assenti contenuti incorporati o collegamenti ipertestuali verso oggetti esterni come le immagini reperibili nel web. In questo caso, qualora il collegamento venisse meno per qualsiasi ragione, l'immagine non sarebbe più visibile e di conseguenza il contenuto originario ne risulterebbe compromesso.

3.7 AUTO-DOCUMENTAZIONE

Un formato è auto-documentato in riferimento alla sua capacità di supportare l'inclusione di metadati.

3.8 INDIPENDENZA DAL DISPOSITIVO

Definita anche portabilità, fa riferimento alla capacità di un file (codificato in un determinato formato), di essere rappresentabile in maniera fedele, indipendentemente dalla piattaforma hardware e software.

4 ELENCO DEI FORMATI AMMESSI DAL SISTEMA

Per ogni formato si riportano le specifiche di riferimento e una nota sull'idoneità o meno dello stesso alla conservazione con relativo workaround.

- C = Conservabile
- G = Gestito - ma non idoneo alla conservazione (vedere NB al [§1](#)).
Una nota identifica quei formati che sono stati deprecati ovvero formati (già conservati) oggi non più idonei alla conservazione

NB:

- Si ricorda che per i formati busta e archivio (es: EML, ZIP, P7M ecc...) il Conservatore non esegue controlli di validazione sui file contenuti; questa verifica, da eseguirsi in fase di formazione dei documenti, rimane a carico del Soggetto Produttore.
- Ogni sistema di gestione documentale deve adattare le strutture proposte al contesto e alla tipologia documentale specifica, oggetto del versamento; in questo senso, significa che per ogni flusso di gestione documentale dovrà essere possibile definire ed identificare i formati adottati nelle varie fasi del ciclo di vita del documento (bozza/formazione >> consolidato/registrazione >> archiviato/gestione >> conservazione).
- Questo elenco può essere integrato secondo le indicazioni del Soggetto Produttore per includere nuovi formati file previsti per i documenti destinati alla conservazione digitale
- Ogni Soggetto Produttore può definire di sotto-insiemi di questi, come elenchi di formati specifici per ogni flusso di versamento previsto dal suo Piano/Manuale di Gestione e Conservazione Documentale.

4.1 Tabella dei formati

Estensione attesa	MIME-TYPE	Conservazione	Note
.3gp	video/3gpp	G	Deprecato dal 01/10/2025
.7z	application/x-7z-compressed	C	Formato archivio
.accdb	application/msaccess	G	Database Microsoft Access
.ai	application/illustrator	G	
.aiff	audio/aiff	G	
.ari	image/arriraw	G	RAW Arri Alexa
.avi	video/x-msvideo	G	Deprecato dal 01/10/2025 (Standard RIFF)
.bin	application/octet-stream	G	Flusso binario generico
.bmp	image/x-ms-bmp	C	
.bz2	application/x-bzip2	C	(Integrazione richiesta)
.css	text/css	C	
.csv	text/csv	C	
.dat	application/vnd.ms-tnef	G	Formato proprietario di Microsoft Outlook/Exchange NON idoneo alla conservazione a lungo termine
.deb	application/x-debian-package	G	Pacchetto software Debian
.dng	image/x-adobe-dng	G	
.doc	application/msword	G	Deprecato dal 01/10/2025
.docx	application/vnd.openxmlformats-officedocument.wordprocessingml.document	G	
.dotx	application/vnd.openxmlformats-officedocument.wordprocessingml.template	G	
.dpx	image/x-dpx	G	
.dtd	application/xml-dtd	C	
.dxf	image/vnd.dxf	G	
.dwf	model/vnd.dwf	G	
.dwg	image/vnd.dwg	C	
.eml	message/rfc822	C	(Mime IETF standard per EML)
.emlx	message/x-emlx	G	Deprecato dal 01/10/2025
.eps	application/postscript	G	
.epub	application/epub+zip	G	Deprecato dal 01/10/2025
.exe	application/x-dosexec	G	Eseguibile Windows
.exr	image/x-exr	C	
.fbx	model/vnd.fbx	C	
.fdf	application/vnd.fdf	C	
.gif	image/gif	G	Deprecato dal 01/10/2025

Estensione attesa	MIME-TYPE	Conservazione	Note
.gz	application/x-compressed	G	Deprecato dal 01/10/2025
.gz	application/x-gzip	G	Deprecato dal 01/10/2025
.html	text/html	C	Conservare insieme al CSS
.idml	application/x-indesign+xml	G	
.jpeg, .jpg	image/jpeg	C	Solo se nativo
.json	application/json	C	Con schema associato
.jsonld	application/ld+json	C	
.log	text/x-log	C	
.mbox	application/mbox	C	
.mdb	application/msaccess	G	(Variante legacy Access)
.md	text/markdown	C	
.mgs	application/vnd.ms-outlook	C	
.mht	message/rfc822	C	MHTML
.mkv	video/x-matroska	C	Formato contenitore aperto (Matroska)
.mml	text/mathml	C	
.mp3	audio/mpeg	G	Deprecato dal 01/10/2025
.mp4	video/mp4	C	
.mpeg	video/mpeg	G	Deprecato dal 01/10/2025
.msi	application/x-ms-installer	G	Pacchetto installazione Windows
.dll, .bat	application/x-msdownload	G	File sistema/script
.mxf	application/mxf	C	
.odb	application/vnd.oasis.opendocument.data base	G	Deprecato dal 01/10/2025
.odf	application/vnd.oasis.opendocument.form ula	C	
.odg	application/vnd.oasis.opendocument.gra phics	G	
.odi	application/vnd.oasis.opendocument.ima ge	C	
.odp	application/vnd.oasis.opendocument.pre sentation	C	
.ole	application/x-tika-msoffice	G	Contenitore OLE2 (Legacy Office)
.ots	application/vnd.oasis.opendocument.spr eadsheet-template	C	Template Foglio di calcolo
.ods	application/vnd.oasis.opendocument.spr eadsheet	C	
.odt	application/vnd.oasis.opendocument.text	C	
.otf	application/x-font-otf	C	
.ott	application/vnd.oasis.opendocument.text -template	C	Template Testo
.p12	application/x-pkcs12	C	Oggetto crittografico
.p7c, .p7b	application/pkcs7-certificates	C	Allegato tecnico (Certificati)
.p7s	application/pkcs7-signature	C	Firma detached
.p7m	application/pkcs7-mime	C	Busta crittografica S/MIME
.pdf	application/pdf	C	Profili PDF/A e PDF/B

Estensione attesa	MIME-TYPE	Conservazione	Note
.pdf	multipart/appldouble	G	(Resource fork Apple)
.png	image/png	C	
.ppsx	application/vnd.openxmlformats-officedocument.presentationml.presentation	G	Slideshow (Mime unificato)
.ppt	application/vnd.ms-powerpoint	G	
.pptx	application/vnd.openxmlformats-officedocument.presentationml.presentation	G	
.ps	application/postscript	G	
.psd	image/x-psd	G	
.pst	application/vnd.ms-outlook	C	
.rar	application/x-rar-compressed	G	Deprecato dal 01/10/2025
.rtf	application/rtf	G	Deprecato dal 01/10/2025
.rtx	text/richtext	C	
.sql	application/sql	C	
.stl	model/stl	G	
.svg	image/svg+xml	C	
.sxw	application/vnd.sun.xml.writer	C	OpenOffice 1.0
.tar	application/x-tar	C	
.tar.gz o .tgz	application/gzip	C	
.tex	application/x-tex	G	
.tiff	image/tiff	C	Senza compressione
.tsd	application/timestamped-data	C	Marca temporale
.ttf	application/x-font-ttf	C	
.txt	text/plain	C	
.wav	audio/x-wav	C	Senza compressione
.webp	image/webp	C	
.wma	audio/x-ms-wma	G	Deprecato dal 01/10/2025
.wmv	video/x-msvideo	G	Deprecato dal 01/10/2025
.woff	application/font-woff	C	
.xhtml	application/xhtml+xml	C	Conservare con CSS
.xls	application/vnd.ms-excel	G	Deprecato dal 01/10/2025
.xslt	application/xslt+xml	G	Deprecato dal 01/10/2025
.xlsx	application/vnd.openxmlformats-officedocument.spreadsheetml.sheet	G	
.xltx	application/vnd.openxmlformats-officedocument.spreadsheetml.template	G	
.xml	application/xml	C	
.dtd	application/xml-dtd	C	
.zip	application/zip	C	
.zip	application/vnd.etsi.asic-s+zip	C	Busta firma ASiC-S
.zip	application/x-tika-ooxml-protected	G	File OOXML protetti da password

5 SPECIFICHE DI RAPPRESENTAZIONE

Viste le sempre più veloci nuove introduzioni ed evoluzioni tecnologiche, i documenti conservati secondo queste specifiche non saranno più abbinati ad un unico specifico software di visualizzazione (viewer) da includere nel pacchetto di distribuzione per l'Esibizione a Norma dei documenti conservati; in alternativa, per ognuno dei formati analizzati, riportiamo indicazione delle specifiche (o del producer) di riferimento, da tenere in considerazione per la corretta rappresentazione del contenuto del file.

5.1 Mime-Type previsti

Estensione attesa	MIME-TYPE	Specifiche
.3gp	video/3gpp	3GPP TS 26.244; ISO/IEC 14496-12
.7z	application/x-7z-compressed	7z Format Description / LZMA SDK
.accdb	application/msaccess	Microsoft Access 2019 Proprietary Format
.ai	application/illustrator	Adobe Illustrator File Format Specification
.aiff	audio/aiff	Apple Audio Interchange File Format (AIFF)
.ari	image/arriraw	SMPTE RDD 30:2014
.avi	video/x-msvideo	Microsoft RIFF; OpenDML AVI File Format
.bin	application/octet-stream	Binary Data Stream (Generic)
.bmp	image/x-ms-bmp	Microsoft Bitmap File Format
.bz2	application/x-bzip2	Bzip2 Compressed Archive Specification
.css	text/css	W3C Recommendation CSS 2.1; CSS3 Modules
.csv	text/csv	RFC 4180: Common Format and MIME Type for CSV
.dat	application/vnd.ms-tnef	Formato proprietario di Microsoft Outlook/Exchange; Per accedere al suo contenuto - Servizi web gratuiti come Winmaildat.com - Su Windows, Winmail Opener - Su macOS, TNEF's Enough
.deb	application/x-debian-package	Debian Package Format (ar + tar + gzip/xz)
.dng	image/x-adobe-dng	Adobe Digital Negative (DNG) Specification
.doc	application/msword	Microsoft [MS-DOC]: Word Binary File Format
.docx	application/vnd.openxmlformats-officedocument.wordprocessingml.document	ISO/IEC 29500-1 (Office Open XML)
.dotx	application/vnd.openxmlformats-officedocument.wordprocessingml.template	ISO/IEC 29500-1 (Office Open XML Template)
.dpx	image/x-dpx	SMPTE ST 268:2014 (Digital Picture Exchange)
.dtd	application/xml-dtd	W3C XML 1.0 (DTD)
.dxf	image/vnd.dxf	Autodesk DXF Reference
.dwf	model/vnd.dwf	Autodesk DWF Specification
.dwg	image/vnd.dwg	Open Design Specification for .dwg
.eml	message/rfc822	RFC 5322: Internet Message Format
.emlx	message/x-emlx	Apple Email Format (Proprietario)

Estensione attesa	MIME-TYPE	Specifiche
.eps	application/postscript	Adobe Encapsulated PostScript (EPS) v3.0
.epub	application/epub+zip	ISO/IEC 30135; EPUB 3.2 (W3C)
.exe	application/x-dosexec	Microsoft PE (Portable Executable) Format
.exr	image/x-exr	OpenEXR File Layout; SMPTE ST 2065-4
.fbx	model/vnd.fbx	Autodesk FBX Files
.fdf	application/vnd.fdf	ISO 32000 (PDF Forms Data Format)
.gif	image/gif	W3C / CompuServe GIF89a Specification
.gz	application/x-compressed	RFC 1952 (GZIP file format specification)
.gz	application/x-gzip	RFC 6713
.html	text/html	W3C Recommendation HTML 5.2
.idml	application/x-indesign+xml	Adobe InDesign Markup Language (IDML)
.jpeg, .jpg	image/jpeg	ISO/IEC 10918-1 (ITU-T Rec. T.81)
.json	application/json	RFC 8259: The JSON Data Interchange Format
.jsonld	application/ld+json	W3C Recommendation JSON-LD 1.0
.log	text/x-log	Plain Text (ASCII / UTF-8)
.mbox	application/mbox	RFC 4155: The application/mbox Media Type
.mdb	application/msaccess	Microsoft Access Binary Format (Legacy)
.md	text/markdown	RFC 7763: The text/markdown Media Type
.mgs	application/vnd.ms-outlook	Microsoft Outlook Item
.mht	message/rfc822	RFC 822 (MIME Encapsulation)
.mkv	video/x-matroska	Matroska Media Container Specification; IETF RFC 8794
.mml	text/mathml	W3C MathML Recommendation
.mp3	audio/mpeg	ISO/IEC 11172-3 (MPEG-1 Audio Layer 3)
.mp4	video/mp4	ISO/IEC 14496-14 (MP4 File Format)
.mpeg	video/mpeg	ISO/IEC 11172-2 (MPEG-1 Video)
.msi	application/x-ms-installer	Microsoft OLE Structured Storage (Installer)
.dll, .bat	application/x-msdownload	Microsoft PE (Portable Executable)
.mxf	application/mxf	SMPTE ST 377 (Material Exchange Format)
.odb	application/vnd.oasis.opendocument.data base	ISO/IEC 26300 (ODF Database)
.odf	application/vnd.oasis.opendocument.form ula	ISO/IEC 26300 (ODF Formula)
.odg	application/vnd.oasis.opendocument.grap hics	ISO/IEC 26300 (ODF Graphics)
.odi	application/vnd.oasis.opendocument.imag e	ISO/IEC 26300 (ODF Image)
.odp	application/vnd.oasis.opendocument.pres entation	ISO/IEC 26300 (ODF Presentation)
.ole	application/x-tika-msoffice	Microsoft OLE 2 Compound Document Format
.ots	application/vnd.oasis.opendocument.spre adsheet-template	ISO/IEC 26300 (ODF Spreadsheet Template)
.ods	application/vnd.oasis.opendocument.spre adsheet	ISO/IEC 26300 (ODF Spreadsheet)
.odt	application/vnd.oasis.opendocument.text	ISO/IEC 26300 (ODF Text)
.otf	application/x-font-otf	ISO/IEC 14496-22 (OpenType)

Estensione attesa	MIME-TYPE	Specifiche
.ott	application/vnd.oasis.opendocument.text-template	ISO/IEC 26300 (ODF Text Template)
.p12	application/x-pkcs12	RFC 7292: PKCS #12
.p7c, .p7b	application/pkcs7-certificates	RFC 8551; RFC 5754
.p7s	application/pkcs7-signature	RFC 2315 (PKCS #7); RFC 5652 (CMS)
.p7m	application/pkcs7-mime	RFC 8551: S/MIME 4.0
.pdf	application/pdf	ISO 32000-1 (PDF 1.7); ISO 19005 (PDF/A)
.pdf	multipart/appledouble	RFC 1740 (AppleDouble format)
.png	image/png	ISO/IEC 15948:2004 (PNG)
.ppsx	application/vnd.openxmlformats-officedocument.presentationml.presentation	ISO/IEC 29500-1 (PresentationML Slideshow)
.ppt	application/vnd.ms-powerpoint	Microsoft [MS-PPT]: PowerPoint Binary File Format
.pptx	application/vnd.openxmlformats-officedocument.presentationml.presentation	ISO/IEC 29500-1 (PresentationML)
.ps	application/postscript	Adobe PostScript Language Reference
.psd	image/x-psd	Adobe Photoshop File Formats Specification
.pst	application/vnd.ms-outlook	Microsoft [MS-PST]: Outlook Personal Folders
.rar	application/x-rar-compressed	RAR File Format Description (rarlab)
.rtf	application/rtf	Microsoft RTF Specification v1.9.1; RFC 1521
.rtx	text/richtext	RFC 1341 (MIME: Richtext)
.sql	application/sql	RFC 6922; ISO/IEC 9075 (SQL)
.stl	model/stl	Stereolithography Interface Format (STL)
.svg	image/svg+xml	W3C Recommendation SVG 1.1
.sxw	application/vnd.sun.xml.writer	OpenOffice.org XML (Legacy)
.tar	application/x-tar	POSIX IEEE P1003.1
.tar.gz	application/gzip	RFC 1951
.tex	application/x-tex	LaTeX Project Public License
.tiff	image/tiff	Adobe TIFF Revision 6.0; ISO 12234-2
.tsd	application/timestamped-data	RFC 3161 (Time-Stamp Protocol)
.ttf	application/x-font-ttf	Apple TrueType Manual; Microsoft Typography
.txt	text/plain	RFC 2046; ISO/IEC 10646 (Unicode)
.wav	audio/x-wav	RFC 2361; Microsoft/IBM WAVE
.webp	image/webp	Google WebP Container Specification
.wma	audio/x-ms-wma	Microsoft WMA Specification
.wmv	video/x-msvideo	Microsoft ASF Specification (WMV)
.woff	application/font-woff	W3C Recommendation WOFF 1.0/2.0
.xhtml	application/xhtml+xml	W3C Recommendation XHTML
.xls	application/vnd.ms-excel	Microsoft [MS-XLS]: Excel Binary File Format
.xslt	application/xslt+xml	W3C XSL Transformations (XSLT)
.xlsx	application/vnd.openxmlformats-officedocument.spreadsheetml.sheet	ISO/IEC 29500-1 (SpreadsheetML)
.xltx	application/vnd.openxmlformats-officedocument.spreadsheetml.template	ISO/IEC 29500-1 (SpreadsheetML Template)
.xml	application/xml	W3C Recommendation XML 1.0

Estensione attesa	MIME-TYPE	Specifiche
.zip	application/zip	ISO/IEC 21320-1; PKWARE APPNOTE
.zip	application/vnd.etsi.asic-s+zip	ETSI TS 102 918 (ASiC-S)
.zip	application/x-tika-ooxml-protected	ISO/IEC 29500 (Encrypted Package)

5.2 Riversamento di formato

Quando uno dei “Formati Conservabili” viene deprecato o va in disuso, il Conservatore può supportare il Soggetto Produttore nelle previste attività di riversamento.

I formati deprecati non dovrebbero essere più utilizzati per produrre file destinati alla conservazione a lungo termine, ma verranno comunque ammessi a Sistema con l’opzione “G”.

Riguardo ai file (G) già conservati, prima che il formato di rappresentazione diventasse obsoleto, il conservatore produrrà una volta all’anno l’elenco delle UD (documenti) e PdA (archivi) candidati per il riversamento di formato.

Al fine di garantirne la leggibilità e intelligibilità, prima che questi diventino definitivamente illeggibili (quindi non più supportati dai software di visualizzazione o per l’obsolescenza dei relativi sistemi operativi) il Produttore dovrà indicare l’opzione di riversamento prescelta. In mancanza di tale indicazione il conservatore non potrà procedere al riversamento di formato.

Le opzioni possibili per il riversamento dipendono dai formati in questione: quando il Soggetto Produttore è una Pubblica Amministrazione (art. 2.2 del CAD), il riversamento di formato deve essere eseguito secondo il processo certificato descritto nell’allegato 3 delle linee guida AgID sulla Formazione, Gestione e Conservazione dei documenti informatici; il Conservatore può supportare il produttore attraverso l’utilizzo di strumenti ritenuti idonei e nel predisporre il verbale di certificazione che il Pubblico Ufficiale, incaricato dal Produttore, dovrà sottoscrivere al fine di validare il processo di riversamento.

È essenziale che i le UD oggetto di riversamento siano poi conservate (riversamento), accompagnate da un verbale/atto di conformità che riporti le impronte hash, prima e dopo la conversione, di ogni file.

5.2.1 Mappatura di riversamento consigliata

Tipologia	Formato Deprecato (Origine)	Formato Target (Destinazione)	Criteri
Office Legacy	.doc, .xls, .ppt, .sxw	PDF/A (.pdf)	Standard ISO 19005, autocontenuto (font inclusi), indipendente dal SW.
Email	.msg (Outlook)	EML (.eml) o PDF/A	RFC 5322 è testo in chiaro e standard aperto; il PDF garantisce la "fotografia" del messaggio.
Audio	.wma, .mp3 (legacy)	WAV (PCM) o MP4 (AAC)	.wav è non compresso e robusto; .mp4 è lo standard ISO/IEC supportato ovunque.
Video	.avi, .wmv, .3gp	MP4 (H.264/AVC)	Standard ISO/IEC 14496-14, ampia diffusione e stabilità.
Immagini	.psd, .psp	TIFF o PNG	Formati lossless (senza perdita), aperti e ben documentati ISO/W3C.
Archivi	.rar, .ace, .lzh	ZIP	Standard ISO 21320-1, ampiamente supportato.

5.2.2 Strumenti Software Consigliati

Si propone qui di seguito un elenco di strumenti utilizzabili in ambito professionale, prestando attenzione alle specifiche licenze.

Scenario A. Documenti Office e Testo (doc, xls, ppt -> pdf, odt)

LibreOffice (motore di conversione)

- Licenza: MPL v2.0 (Mozilla Public License).
- È il gold-standard open source per interpretare formati Microsoft legacy (.doc binari) e convertirli in PDF/A-1b o PDF/A-2b.

Scenario B. Immagini e Grafica (psd, bmp, tiff proprietari -> tiff, png)

ImageMagick

- Licenza: Apache 2.0.
- Gestisce oltre 200 formati immagine. Permette di "appiattare" livelli (flatten) da PSD o convertire RAW proprietari in TIFF standard non compressi.

Scenario C. Audio e Video (avi, wmv, wma -> mp4, wav)

FFmpeg

- Licenza: LGPL v2.1+ (GPL opzionale). Alcuni codec (es. AAC o H.264) possono essere coperti da brevetti (MPEG-LA).
- **Alternative:** Convertire in formati video MKV (contenitore) con codec video FFV1 (lossless) e audio FLAC o PCM. Per restare su MP4, si usa FFmpeg con codec libx264.

Scenario D. Email (.msg -> .eml)

MsgConvert (parte di Email::Outlook::Message)

- Linguaggio: Perl
- Licenza: Open Source (Artistic License / GPL).
- Converte fedelmente i file binari proprietari Outlook (.msg) nel formato testo standard RFC 5322 (.eml), conservando header e allegati.