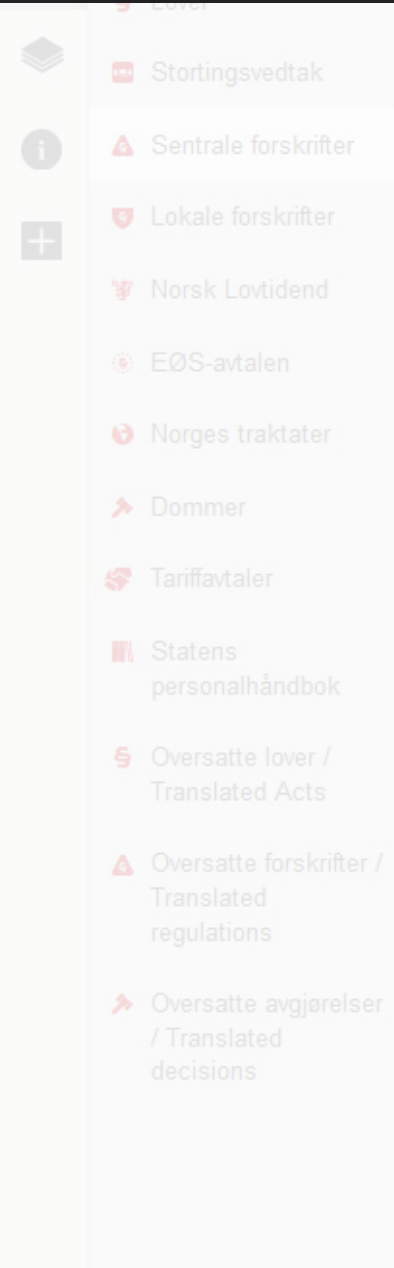




# Digitale filformater

Forvaltningsforum – 20200528



0 Endret ved forskrift 26 nov 2019 nr. 1582 (i kraft 1 jan 2020).

## Del III. Formatkrav m.m. ved uttak fra databasetabeller

### § 5-10. Organisering av arkivuttrekk

(1) Arkivuttrekk fra databaser skal bestå av sekvensielle tekstfiler. Et uttrekk kan bestå av en eller flere filer. Relasjoner mellom filene i uttrekket skal være beskrevet i den medfølgende dokumentasjonen slik det fremgår av § 5-24.

(2) Data skal tas ut på en måte som gjør at hver posttype eksporteres til én enkelt sekvensiell fil. Opprinnelige fysiske pekere skal omgjøres til unike logiske nøkler (primær- og fremmednøkler). Felt som er repetert flere ganger innenfor en post, skal skilles ut fra den opprinnelige posten, og eksporteres som egne poster i en ny fil med tilhørende nøkler. Pakkede og binære tallfelt aksepteres ikke i arkivuttrekk ved avlevering og deponering. For øvrig henvises det til § 5-11 til § 5-20.

(3) For arkivuttrekk i XML-format kan ovennevnte krav avvikles etter avtale med Arkiverket.

(4) For Noark-systemer gjelder § 5-21 til § 5-23.

### § 5-11. Tegnssett

(1) Arkivuttrekk og medfølgende struktur- og innholdsbeskrivelser skal overføres som ren tekst i ukryptert form, og benytte godkjent tegnssett.

(2) Godkjente tegnssett er:

- Unicode UTF-8 (ISO/IEC 10646-1:2000 Annex D)
- ISO 8859-1:1998, Latin 1
- ISO 8859-4:1998, Latin-4 for samiske tegn.

(3) Andre tegnssett aksepteres bare etter avtale med Arkiverket.

### § 5-12. Tekstfilformater i arkivuttrekk

Tekstfiler i et arkivuttrekk skal ha ett av følgende formater:

- Fast format: Alle felt kommer i faste posisjoner innen postene. Lengde og/eller startposisjon for hvert felt skal spesifiseres i struktur- og innholdsbeskrivelsen, som skal følge med uttrekket slik det fremgår av § 5-24. Alle lengder og posisjoner angis ut fra antall tegn.
- Tegnseparert format (kommaseparert format) i henhold til standarden IETF RFC 4180: Poster og felt har variabel lengde og er atskilt med faste, entydige skilletegn. Postskille er påkrevd og skal angis i den medfølgende struktur- og

# Hva skjer?

## Forskriften skal revideres

# Ønske om forenkling av prosess ved endringer

## § 5-17. Godkjente filformater for arkivdokumenter ved avlevering eller deponering

I gjeldende § 5-17 listes det opp hvilke filformater som til en hver tid er gyldige for overføring av arkivdokumenter til depot. Behovet for å oppdatere og gjøre endringer i hva som er gyldige filformater skjer i takt med utviklingen av internasjonale krav og standarder og standardiseringsmiljø i Norge. Disse endringene skjer hyppigere enn endringer i RAF legger opp til.

For å imøtekomme behov for oppdatering i takt med den generelle utviklingen vil det være mer hensiktsmessig å løfte ut oppramsingen av enkeltformatene og heller gi en generell beskrivelse av hvilke kriterier et gyldig filformat skal inneholde.

Selve oversikten over hvilke filformater som er gyldige kunngjøres av Riksarkivaren (hjemmeside).

Slike universelle kriterier ble beskrevet i SAMDOK-rapporten fra 2017- og kan fungere som et utgangspunkt for ny tekst til §5-17:

[https://samdokdotcom.files.wordpress.com/2017/12/sluttrapport-2017-samdok-eforv\\_2-filformatsforvaltning-20171025\\_1.pdf](https://samdokdotcom.files.wordpress.com/2017/12/sluttrapport-2017-samdok-eforv_2-filformatsforvaltning-20171025_1.pdf)

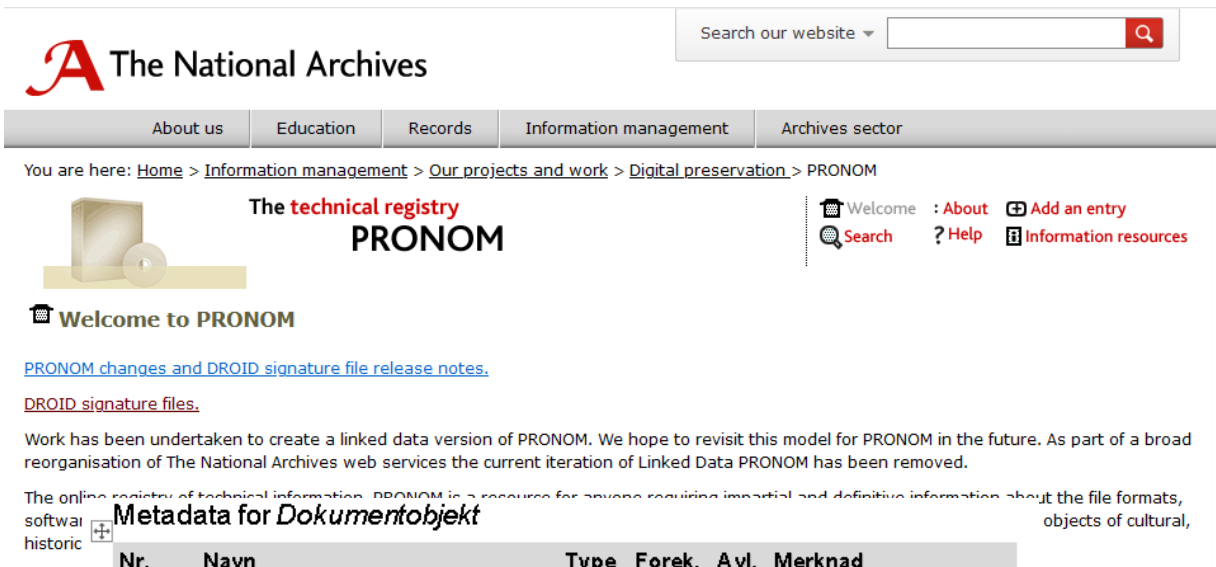
Riksarkivaren fastsetter hvilke filformater som er godkjente ved avlevering eller deponering. Filformatene skal oppfylle kriterier:

- Formatet skal følge en internasjonal standard (ISO, ANSI eller lign.)
- Formatet skal være åpent spesifisert
- Formatet skal være i utstrakt bruk
- Formatet skal ikke være lisensbelagt
- Formatet skal ha en lang forventet levetid
- Formatet skal ikke inneholde eksterne avhengigheter
- Formatet skal være uavhengig av operativsystem og presentasjonsplattform

	ukryptert form, og benytte godkjent tegnsett	
§§ 5-17, 5-18, 5-19 og 5-20	Godkjente filformater for arkivdokumenter ved avlevering eller deponering. Følgende formater er godkjente ved avlevering eller deponering a-j	Vurdere å skille ut godkjente filformater til våre hjemmesider. Filformater endres raskere enn vi klarer å oppdatere forskriften.

**NB: Interne og foreløpige notater**

# Filformat standarder



The National Archives

Search our website

About us | Education | Records | Information management | Archives sector

You are here: Home > Information management > Our projects and work > Digital preservation > PRONOM

**The technical registry**  
**PRONOM**

Welcome to PRONOM

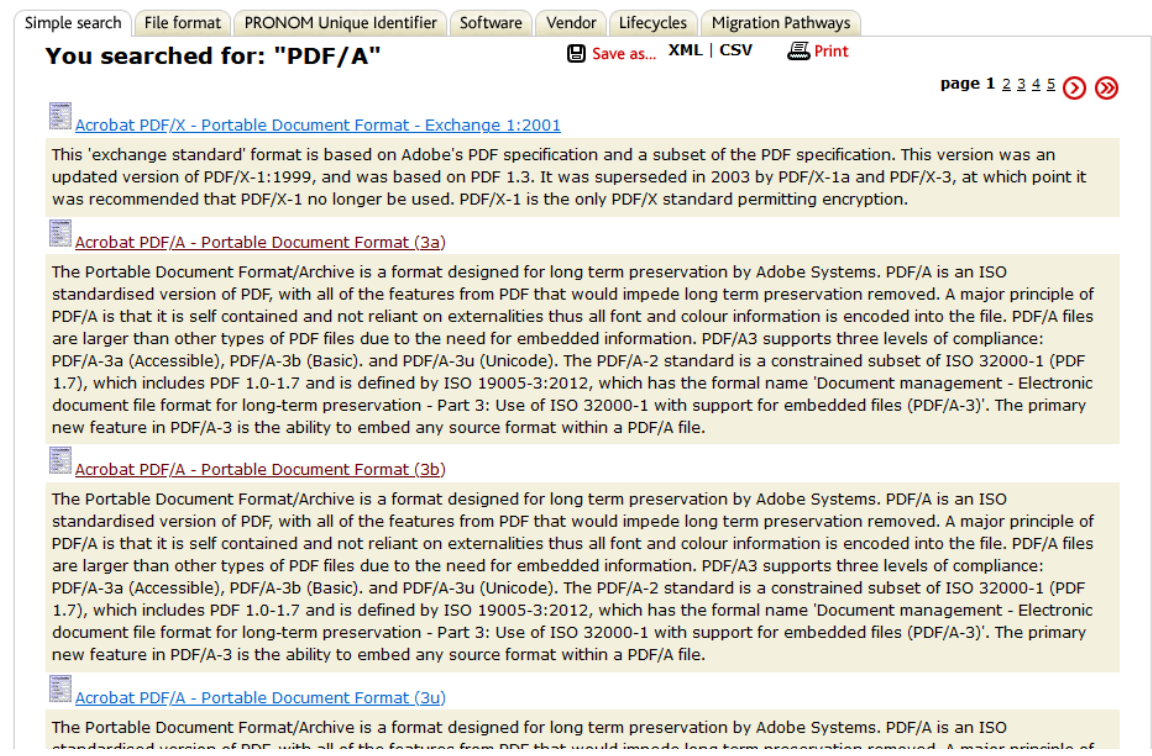
PRONOM changes and DROID signature file release notes.  
DROID signature files.

Work has been undertaken to create a linked data version of PRONOM. We hope to revisit this model for PRONOM in the future. As part of a broad reorganisation of The National Archives web services the current iteration of Linked Data PRONOM has been removed.

The online registry of technical information, PRONOM is a resource for anyone requiring impartial and definitive information about the file formats, objects of cultural, historic

**Metadata for Dokumentobjekt**

Nr.	Navn	Type	Forek.	Avl.	Merknad
M005	versjonsnummer	○	En	A	Identifikasjon av versjoner innenfor samme dokument. Se også merknad 1 nedenfor.
M700	variantformat	○	En	A	Angivelse av om det dreier seg om produksjonsformat, arkivformat, skjermet dokument osv. Se også merknad 2 nedenfor.
M701	format	○	En	A	
M702	formatDetalier	○	En	A	
M600	opprinnelse	○	En	A	



Simple search | File format | PRONOM Unique Identifier | Software | Vendor | Lifecycles | Migration Pathways

You searched for: "PDF/A"

Save as... | XML | CSV | Print

page 1 2 3 4 5

**Acrobat PDF/X - Portable Document Format - Exchange 1:2001**

This 'exchange standard' format is based on Adobe's PDF specification and a subset of the PDF specification. This version was an updated version of PDF/X-1:1999, and was based on PDF 1.3. It was superseded in 2003 by PDF/X-1a and PDF/X-3, at which point it was recommended that PDF/X-1 no longer be used. PDF/X-1 is the only PDF/X standard permitting encryption.

**Acrobat PDF/A - Portable Document Format (3a)**

The Portable Document Format/Archive is a format designed for long term preservation by Adobe Systems. PDF/A is an ISO standardised version of PDF, with all of the features from PDF that would impede long term preservation removed. A major principle of PDF/A is that it is self contained and not reliant on externalities thus all font and colour information is encoded into the file. PDF/A files are larger than other types of PDF files due to the need for embedded information. PDF/A3 supports three levels of compliance: PDF/A-3a (Accessible), PDF/A-3b (Basic), and PDF/A-3u (Unicode). The PDF/A-2 standard is a constrained subset of ISO 32000-1 (PDF 1.7), which includes PDF 1.0-1.7 and is defined by ISO 19005-3:2012, which has the formal name 'Document management - Electronic document file format for long-term preservation - Part 3: Use of ISO 32000-1 with support for embedded files (PDF/A-3)'. The primary new feature in PDF/A-3 is the ability to embed any source format within a PDF/A file.

**Acrobat PDF/A - Portable Document Format (3b)**

The Portable Document Format/Archive is a format designed for long term preservation by Adobe Systems. PDF/A is an ISO standardised version of PDF, with all of the features from PDF that would impede long term preservation removed. A major principle of PDF/A is that it is self contained and not reliant on externalities thus all font and colour information is encoded into the file. PDF/A files are larger than other types of PDF files due to the need for embedded information. PDF/A3 supports three levels of compliance: PDF/A-3a (Accessible), PDF/A-3b (Basic), and PDF/A-3u (Unicode). The PDF/A-2 standard is a constrained subset of ISO 32000-1 (PDF 1.7), which includes PDF 1.0-1.7 and is defined by ISO 19005-3:2012, which has the formal name 'Document management - Electronic document file format for long-term preservation - Part 3: Use of ISO 32000-1 with support for embedded files (PDF/A-3)'. The primary new feature in PDF/A-3 is the ability to embed any source format within a PDF/A file.

**Acrobat PDF/A - Portable Document Format (3u)**

The Portable Document Format/Archive is a format designed for long term preservation by Adobe Systems. PDF/A is an ISO standardised version of PDF, with all of the features from PDF that would impede long term preservation removed. A major principle of

# PRONOM mapping vil bli en del av informasjon tilhørende Forskriften

## § 5-17. Godkjente filformater for arkivdokumenter ved avlevering eller deponering

(1) Følgende formater er godkjente ved avlevering eller deponering:

a) Arkivdokumenter med ren tekst:

- TXT<sup>45</sup> (*x-fmt/111, x-fmt/16*) – ren tekstfil uten innkapslede objekter, ref. spesifikasjon i § 5-11.
- TIFF<sup>67</sup> (*se § 5-18*); – Tag Image File Format versjon 6, avbildninger av tekst, med de presiseringer som fremgår av § 5-18, første ledd.
- PDF/A (*se § 5-19*) – med de presiseringer som fremgår av § 5-18, andre ledd.
- XML (*se § 5-19*) – Extensible Markup Language versjon 1.0, med de presiseringer som fremgår av § 5-19.
- For tekstdokumenter i OOXML eller ODF format bør en PDF/A-versjon leveres som tillegg.

b) Arkivdokumenter inneholdende tekst med objekter:

- TIFF (*se § 5-18*) – Tag Image File Format versjon 6, avbildninger, med de presiseringer som fremgår av § 5-18, første ledd.
- PDF/A (*se § 5-19*) – med de presiseringer som fremgår av § 5-18, andre ledd.

c) For digitale fotografier og bilder aksepteres følgende formater:

- TIFF (*se § 5-18*) versjon 6 med de presiseringer som fremgår av § 5-18, første ledd.
- JPEG<sup>8</sup> (*fmt/42-44*) (ISO/IEC 10918-1:1994).
- JPEG 2000 (*x-fmt/382*) (ISO/IEC 15444-1).
- PNG<sup>9</sup> (*fmt/11-13*) (ISO/IEC 15948).
- PDF/A (*se § 5-19*) – med de presiseringer som fremgår av § 5-18, andre ledd.

d) For kart aksepteres følgende formater:

- TIFF (*se § 5-18*) versjon 6 med de presiseringer som fremgår av § 5-18, første ledd.
- SOSI<sup>10</sup> (*fmt/1246-1250<sup>11</sup>*) versjon 2.2 (1995) eller nyere.
- GML<sup>12</sup> (*x-fmt/382*) (ISO 19136).

e) For videosekvenser aksepteres følgende formater:

- MPEG-2<sup>13</sup> (*fmt/640, fmt/386*) (ISO/IEC 13818-2).
- MPEG-4/H.264Part 10<sup>14</sup> (*fmt/199*): Advanced Video Coding (ISO/IEC 14496-10).

f) For lydsekvenser aksepteres følgende formater:

- MP3<sup>15</sup> (*fmt/134*) (ISO/IEC 13818-3).
- PCM-basert Wave<sup>16</sup> (*fmt/6, fmt/141, fmt/142*) (ref. Multimedia Programming Interface and Data Specification 1.0 – IBM Corporation og Microsoft Corporation, 1991).
- FLAC<sup>17</sup> (*fmt/659*), som spesifisert på xiph.org.

g) For regneark aksepteres følgende formater:

- PDF/A (*se § 5-19*) – med de presiseringer som fremgår av § 5-18, andre ledd.
- XML (*se § 5-19*), med de presiseringer som fremgår av § 5-19.
- For regneark i OOXML eller ODF format (*se § 5-19*) bør en PDF/A-versjon leveres som tillegg.

h) For web-sider aksepteres følgende formater:

- WARC<sup>18</sup> (*x-fmt/265*) (ISO 28500).
- HTML<sup>19</sup> (*fmt/96-100, fmt/471*) (ISO/IEC 15445) basert på tegnssett som spesifisert i § 5-11.

**Rimer ikke med § 5-20**

# Ønsker om nye formater, endringer, tillegg. Klipp fra noen eposter

NUUG skrev litt om epost - - -

- - - referanser til IETF-standardene for epostformatene.

De er allerede i PRONOM som

<URL: <http://www.nationalarchives.gov.uk/PRONOM/fmt/278> > og

<URL: <http://www.nationalarchives.gov.uk/PRONOM/fmt/950> >.

GIF er beskrevet i <URL: <https://en.wikipedia.org/wiki/GIF> > og er et mye brukt format

- - - flere formater jeg godt kunne godtas som arkivformater,

- \* <https://en.wikipedia.org/wiki/EPUB>
- \* [https://en.wikipedia.org/wiki/Opus\\_\(audio\\_format\)](https://en.wikipedia.org/wiki/Opus_(audio_format))
- \* <https://en.wikipedia.org/wiki/WebM>
- \* <https://en.wikipedia.org/wiki/WebP>
- \* <https://en.wikipedia.org/wiki/Theora>
- \* <https://en.wikipedia.org/wiki/Vorbis>

du, json, noen grunn til at det ikke er et godkjent arkivformat? Annet enn "sånn har det alltid vært"?

Her er den siste «utskrift» vi fikk fra IKT over filtypene vi nå har i ephorte, som en illustrasjon på utfordringen vår

Extension	Col...	Description	> Bytes	% Bytes	Files
.pdf		PDF File	677,1 GB	56,4%	971 8...
.zip		Compressed (zipped) Folder	145,5 GB	12,1%	2 646
.mpg		MPG File	103,8 GB	8,7%	309
.smc		SMC File	98,2 GB	8,2%	645
.jpg		JPEG image	33,7 GB	2,8%	39 948
.docx		Office Open XML Document	21,4 GB	1,8%	75 578
.pptx		PPTX File	18,7 GB	1,6%	2 649
.flc		IFC File	15,8 GB	1,3%	588
.mp4		MP4 File	15,1 GB	1,3%	93
.xml		XML Document	13,2 GB	1,1%	14 163
.gab		GAB File	11,5 GB	1,0%	1 540
.htm		HTML Document	7,4 GB	0,6%	153 4...
.dwg		DWG File	5,7 GB	0,5%	4 730
.msg		MSG File	5,7 GB	0,5%	5 102
.xlsx		XLSX File	5,5 GB	0,5%	20 674
.mov		MOV File	3,3 GB	0,3%	48
.tif		TIFF image	3,0 GB	0,3%	532
.mpp		MPP File	2,7 GB	0,2%	266
.rvt		RVT File	2,6 GB	0,2%	22
.doc		DOC File	2,5 GB	0,2%	7 567
.html		HTML Document	892,4 MB	0,1%	17 785
.png		PNG image	823,4 MB	0,1%	9 224
.xls		XLS File	729,2 MB	0,1%	2 523
.avi		AVI File	693,2 MB	0,1%	12
.dotx		DOTX File	564,1 MB	0,0%	3 455
.ppt		PPT File	562,1 MB	0,0%	81
.bmp		Bitmap image	402,8 MB	0,0%	62
.mpeg		MPEG File	266,0 MB	0,0%	3
.jpeg		JPEG image	245,3 MB	0,0%	344
.partial		Partial Download	212,3 MB	0,0%	6
.vob		VOB File	209,2 MB	0,0%	1
.exe		Application	182,7 MB	0,0%	21
.dxf		DXF File	167,2 MB	0,0%	38
.sdo		SDO File	142,8 MB	0,0%	1 076
.pla		PLA File	142,7 MB	0,0%	2
.rtf		Rich Text Document	137,0 MB	0,0%	1 282
.gal		GAL File	130,4 MB	0,0%	13
.rar		RAR File	124,7 MB	0,0%	21
.pst		PST File	113,6 MB	0,0%	1
.xslm		XLSM File	108,2 MB	0,0%	605
.potx		POTX File	102,2 MB	0,0%	12
.bcfzip		BCFZIP File	87,8 MB	0,0%	5
.7z		7Z File	74,2 MB	0,0%	7
.skp		SKP File	68,0 MB	0,0%	11
.ifczip		IFCZIP File	60,3 MB	0,0%	5
.dat		DAT File	51,6 MB	0,0%	48
.dcm		DCM File	49,9 MB	0,0%	15
.imp.pdf		File	45,2 MB	0,0%	79
		Application extension	43,6 MB	0,0%	49
		FLW File	40,4 MB	0,0%	1
		ZIPX File	40,4 MB	0,0%	1
.xltx		XLTX File	39,0 MB	0,0%	1
.sos		SOS File	35,3 MB	0,0%	1
.gif		GIF image	28,3 MB	0,0%	1
.3dm		3DM File	27,3 MB	0,0%	
.tiff		TIFF image	25,7 MB	0,0%	
.mp2		MP2 File	24,2 MB	0,0%	
.txt		Text Document	17,8 MB	0,0%	11

Frustasjon



# Med referanse til andre også - - -

## Kino og annen (høykvalitets) film

Filformatet avleveres i åpne og standardiserte format, så lenge program og produksjonsutstyr støtter dette. Fortrinnsvis avleveres dette i ett eller flere format for kino-, TV-, VOD- og UHD-master:

- EXR (tapsfri)
- DPX (tapsfri)
- TIFF (tapsfri)
- Jpeg2000 (tapsfri)
- Prores4444
- DNxHR444

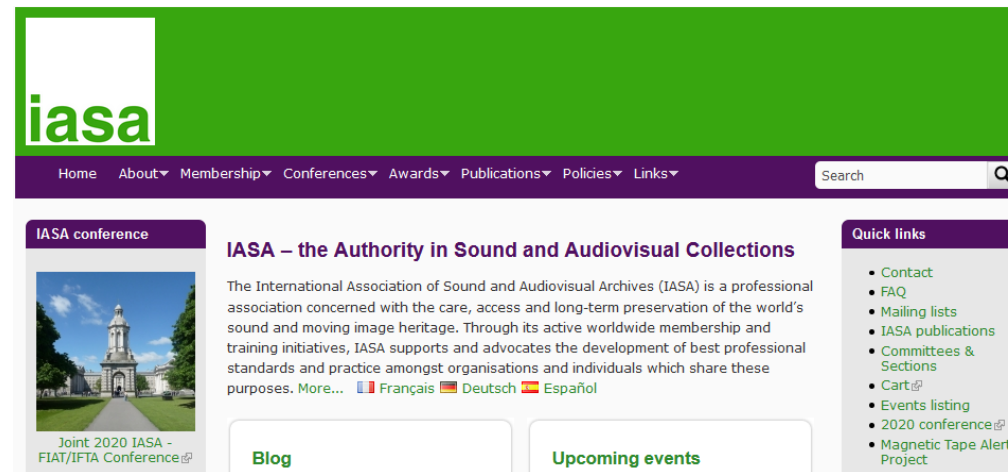
Minimum kvalitet for TV-master og festivalfilm:

- ProresHQ
- DNxHR HQX / HQ

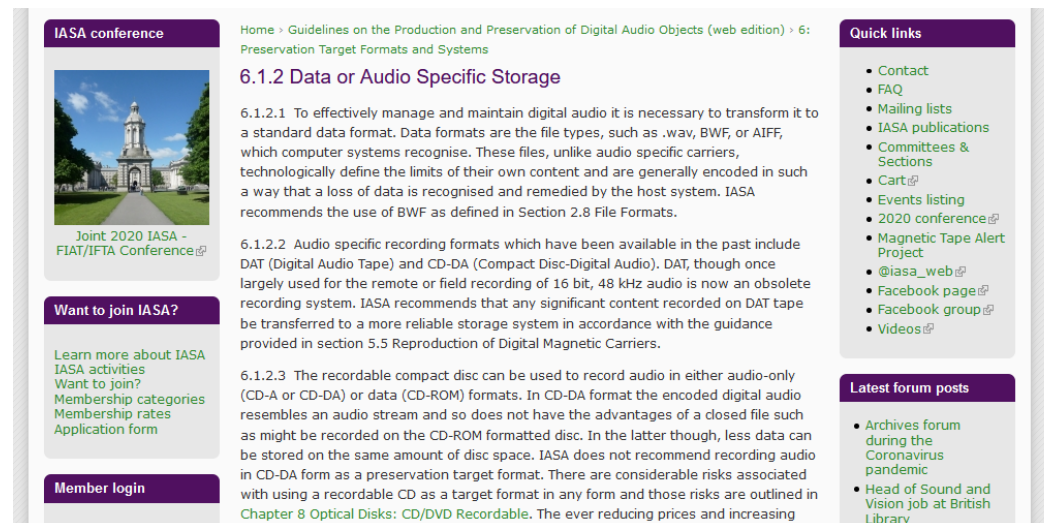
Masterformater oftest brukt for video-, web- og påsynskopier:

- Mpeg4 - H264 (med minimum bitrate 1500 kbit/s)
- Mpeg2 filer (minimum bitrate 50 Mbit/s, og bare I-frame kodet)

(Fra Nasjonalbiblioteket)



The screenshot shows the IASA website homepage. The header is green with the IASA logo. Below the header is a purple navigation bar with links: Home, About, Membership, Conferences, Awards, Publications, Policies, Links. A search bar is on the right. The main content area has a purple sidebar on the left with a 'IASA conference' section featuring a photo of a building and the text 'Joint 2020 IASA - FIAT/IFTA Conference'. The main content area has a purple header 'IASA – the Authority in Sound and Audiovisual Collections' followed by a paragraph of text and a 'More...' link. Below this are two white boxes: 'Blog' and 'Upcoming events'. On the right is a 'Quick links' sidebar with a list of links: Contact, FAQ, Mailing lists, IASA publications, Committees & Sections, Cart, Events listing, 2020 conference, and Magnetic Tape Alert Project.



The screenshot shows a page from the IASA website titled 'Guidelines on the Production and Preservation of Digital Audio Objects (web edition) > 6: Preservation Target Formats and Systems'. The page has a purple sidebar on the left with a 'IASA conference' section featuring a photo of a building and the text 'Joint 2020 IASA - FIAT/IFTA Conference'. Below this is a 'Want to join IASA?' section with links: Learn more about IASA, IASA activities, Want to join?, Membership categories, Membership rates, and Application form. At the bottom is a 'Member login' section. The main content area has a purple header '6.1.2 Data or Audio Specific Storage' followed by three sub-sections: 6.1.2.1, 6.1.2.2, and 6.1.2.3. On the right is a 'Quick links' sidebar with a list of links: Contact, FAQ, Mailing lists, IASA publications, Committees & Sections, Cart, Events listing, 2020 conference, Magnetic Tape Alert Project, @iasa\_web, Facebook page, Facebook group, and Videos. At the bottom is a 'Latest forum posts' sidebar with a list of links: Archives forum during the Coronavirus pandemic and Head of Sound and Vision job at British Library.

Et par andre problemstillinger som kom opp - -

Overskrifter i et notat tilknyttet dette prosjektet:

### 3 Krav ved testing - nasjonale tillegg vs. internasjonale standarder

I Arkivverkets testprosedyrer nå inngår det nå at arkivuttrekk er i henhold til de angitte – vanligvis internasjonale standarder (ISO etc.).

### 5 «Inni eller utafor»

Skal metadata embeddes i filene – eller vedlegges separat? |



# Jotne IT

[JOTNE INFORMATION TECHNOLOGY](#) [SOLUTIONS](#) [PRODUCTS](#) [NEWS & EVENTS](#) [ABOUT JOTNE](#)

Etableringsmøte for «Langtidsarkivering av informasjonsmodeller (digital tvilling/BIM/IFC)»



# Ønsker om tilbakemelding

1. Kommentarer til Forskriftens utforming
2. Ønsker om endringer/tillegg til Forskriften  
- spesielt: mangler vi noe
3. Hvor strenge bør man være – når unntak skal innvilges
4. Innspill til PRONOM prosess