

Formatpolicy för långsiktigt digitalt bevarande

DOKUMENTHISTORIK

Dokumentversion	Datum	Kommentar	Författare
1.00	2021-11-16		Theo Erbenius

Detta är ett arbetsdokument. Om/när så bedöms nödvändigt kan nya kategorier och format komma att införas. Även formatregler och bevarandestrategier kan komma att uppdateras över tid om så bedöms relevant utifrån nya kunskaper eller teknologiska framsteg inom digitalt bevarande.

TERMINOLOGI

- Leveransformat: Det format som det digitala objektet var sparat i när det levererades till Folkrorelsearkivet för Uppsala län.
- Bevarandekopia: En kopia som tagits av en digital fil i syfte att överföra filens innehåll till ett format som är dugligt för digitalt långtidsbevarande.
- Accesskopia: En kopia som skapats av en digital fil i syfte att enkelt kunna tillgängliggöra informationen om/när så är önskvärt.
- Migrering: Överföring av filens innehåll från ett filformat till ett annat (görs vanligtvis i syfte att skapa nya bevarande och/eller accesskopior).
- Föredragna bevarandeformat: Format som bedöms vara både långsiktigt hållbart och som kan agera som migrationsmål för andra format.
- Accepterade bevarandeformat:
 - Kategori 1: Dessa format kan användas för långsiktigt bevarande såväl som utgöra målformat vid migrering i de fall inga föredragna bevarandeformat existerar.
 - Kategori 2: Dessa format kan användas för långsiktigt digitalt bevarande, men inte som målformat vid migrering mellan olika format.
- Normaliseringsverktyg: De program som används i syfte att migrera/överföra filinnehåll från ett filformat till ett annat, vanligtvis i syfte att skapa nya bevarande- och/eller accesskopior.
- Arkivpaket (AIP): De paket vari digitala filer förvaras långsiktigt inom vår e-arkivlösning.

KÄLLOR TILL DETTA DOKUMENT

- https://kost-ceco.ch/cms/publications_de.html
- <https://www.loc.gov/preservation/resources/rfs/>
- <https://digitalpreservation.fi/en/specifications>
- https://wiki.dpconline.org/index.php?title=File_Formats_Assessments
- <https://guides.library.cornell.edu/ecommons/formats>
- <https://riksarkivet.se/rafs?pdf=rafs/RA-FS%202009-02.pdf>

REGELVERK

- I de arkivpaket som skapas under e-arkiveringsprocessen förvaras alltid insända filer i samma format som de var i när de levererades till Folkrorelsearkivet för Uppsala län.
- I de fall det bedöms nödvändigt skapar Folkrorelsearkivet för Uppsala län bevarandekopior (under förutsättning att det är praktiskt genomförbart) av insända filer. Även dessa bevarandekopior sparas i tillhörande arkivpaket.
- I de fall det inlevererade materialet skall/kan tillgängliggöras via e-arkivets tillhörande visningsyta, <https://enskildaarkivet.se/index.php/folkrorelsearkivet-for-uppsala-lan>, tas vid behov accesskopior av de inlevererade filerna.
- I de fall ingen bevarandekopia tas i samband med e-arkiveringsprocessen är det p.g.a. någon av följande anledningar:
 - Den inlämnade filen är redan i sparat i ett format som bedöms långsiktigt hållbart.
 - Det bedöms vara riskfritt att utföra konverteringar till specifika bevarandeformat först vid behov. Detta gäller för närvarande enbart de bild-, ljud- och videoformat som kan hanteras av normaliseringsverktyget FFMPEG eller ImageMagick. Den bakomliggande logiken är att så länge det finns etablerade, licensfria och säkra metoder för att utföra förlustfria konverteringar så är bevarandet av filernas innehåll inte avhängigt vilket format som för stunden utgör dess "behållare".
 - Det finns inga tillgängliga normaliseringsverktyg och/eller dugliga bevarandeformat varigenom/vartill filens innehåll kan överföras.

Digital objektkategori	Leveransformat	Föredragna bevarandeformat	Accepterade bevarandeformat	Bevarandekopia	Accesskopia	Normaliseringsverktyg
Bild (bitmap/raster)	TIFF Baseline 6.0 upp till <1 GB. Tillägg: LZW komprimering och multipage tillåts.	Ja	Nej	Nej	JPEG	ImageMagick
	TIFF 6.0; TIFF/EP; TIFF/IIT; TIFF/FX	Nej	Kategori 2	Nej	JPEG	ImageMagick
	JPEG	Nej	Kategori 2	Nej	JPEG	ImageMagick
	JP2 (JPEG 2000 part 1)	Nej	Kategori 2	Nej	JPEG	ImageMagick
	PNG	Ja	Nej	Nej	JPEG	ImageMagick
	GIF	Nej	Kategori 2	Nej	Nej	ImageMagick
	PDF/A-2 med JPEG 2000 komprimerat bildinnehåll. ***** Lämpligt format för digitiserade textdokument	Ja	Nej	Nej	Nej	Nej
Bild (raw)	ARW, CR2, CRW, DCR, MRW, NEF, ORF, PEF, RAF, X3F	Nej	Nej	Nej	JPEG	ImageMagick
	DNG	Nej	Kategori 1	Nej	JPEG	ImageMagick
Text	DOC	Nej	Nej	PDF/A-2u	PDF/A-2u	Under utredning
	DOCX	Nej	Nej	PDF/A-2u	PDF/A-2u	Under utredning
	ODT, ODF	Nej	Nej	PDF/A-2u	PDF/A-2u	Under utredning
	RTF	Nej	Nej	PDF/A-2u	PDF/A-2u	Under utredning
	WPD	Nej	Nej	PDF/A-2u	PDF/A-2u	Under utredning
	PDF	Nej	Nej	PDF/A-2u	PDF/A-2u	Under utredning
	TXT	Ja	Nej	Nej	Nej	Nej
Kalkylblad	CSV	Nej	Kategori 2	Nej	Nej	Nej
	TAB	Nej	Kategori 2	Nej	Nej	Nej
	XLS	Nej	Kategori 2	Nej	Nej	Nej
	XLSX	Nej	Kategori 1	Nej	Nej	Under utredning
	ODS	Nej	Kategori 1	Nej	Nej	Under utredning
Presentationer	PDF/A-2u	Ja	Nej	Nej	Nej	Under utredning
	PPT	Nej	Nej	PDF/A-2u	PDF/A-2u	Under utredning
	PPTX	Nej	Nej	PDF/A-2u	PDF/A-2u	Under utredning
	ODP	Nej	Nej	PDF/A-2u	PDF/A-2u	Under utredning
Ljud	PCM WAVE	Ja	Nej	Nej	MP3	FFMPEG
	Broadcast WAVE	Nej	Kategori 2	Nej	MP3	FFMPEG
	AIFF	Nej	Kategori 2	Nej	MP3	FFMPEG
	FLAC	Ja	Nej	Nej	MP3	FFMPEG
	MP3	Nej	Kategori 2	Nej	Nej	FFMPEG
	MP4, MP4A, AAC	Nej	Kategori 2	Nej	MP3	FFMPEG
Video	MKV/FFV1	Ja	Nej	Nej	MPEG-4	FFMPEG
	AVI	Nej	Kategori 2	Nej	MPEG-4	FFMPEG
	MOV	Nej	Kategori 2	Nej	MPEG-4	FFMPEG
	MPEG-2	Nej	Kategori 2	Nej	MPEG-4	FFMPEG
	MPEG-4	Nej	Kategori 2	Nej	Nej	FFMPEG
	Motion JPEG 2000	Nej	Kategori 2	Nej	MPEG-4	FFMPEG
Databaser	CSV	Nej	Kategori 2	Nej	Nej	Nej
	TAB	Nej	Kategori 2	Nej	Nej	Nej
	SIARD	Ja	Nej	Nej	Nej	Under utredning
	MS Access 2007 or newer DB/2 or newer H2 database 1.4 or newer MySQL (or MariaDB) 5.5 or newer Oracle 10 or newer PostgreSQL 11 or newer SQL Server 2012 or newer	Nej	Nej	SIARD	Nej	Nej
Webbarkiv	WARC	Nej	Kategori 2	Nej	Nej	Nej