



Værktøjer til validering af PDF(A) og AFP(A)

Kort kig oppefra og info til at arbejde videre med

Jan Hougaard Rasmussen, 9. maj 2016





Table of contents

- 3 - Den Danske lov
- 8 - PDF/A archiving standard
- 11 - AFP/A archiving standard
- 12 - PDF/A Check
- 14 - AFP/A Check

Den Danske lov

- Datatilsynet
- Skat
- Digital bevaring - <http://digitalbevaring.dk/>



Datatilsynet www.datatilsynet.dk

Datatilsynet er den centrale uafhængige myndighed, der fører tilsyn med, at reglerne i persondataloven overholdes. Datatilsynet består af et råd - Datarådet - og et sekretariat.

Datatilsynet bl.a. rådgiver og vejleder, behandler klager og gennemfører inspektioner hos myndigheder og virksomheder.

INGEN DOKUMENTER OMKRING ARKIVERING & FORMAT



Skat <http://www.skat.dk/>

Dokumenter på SKAT beskriver elektronisk arkivering som godkendt "papier".

INGEN DOKUMENTER OMKRING arkiverings"sprog" & FORMAT



Digital bevaring - <http://digitalbevaring.dk/>



Arkivloven er det overordnede regelgrundlag for arkivvæsenet i Danmark.

Arkivloven regulerer de offentlige arkivers arbejde, herunder statslige myndigheders aflevering til offentligt arkiv og brugeres adgang til materiale, under hensyn til tilgængelighedsfrister.

Bevaringsværdige digitale arkivalier skal afleveres i form af arkiveringsversioner. En arkiveringsversion er en kopi af data i et [standardiseret](#), systemuafhængigt [format](#), med bestemte regler for tegnsæt, dokumentformater, struktur og dokumentation. De nærmere specifikationer for arkiveringsversioner er udstedt af Rigsarkivet i [bekendtgørelsesform](#). Vejledninger og hjælpematerialer til bekendtgørelsen findes på [Rigsarkivets hjemmeside](#).

Digital bevaring - <http://digitalbevaring.dk/>

Udvalgte formater og deres egenskaber

Der er ikke nogen entydig formel for hvilke formater man skal vælge til bevaring. Det vil altid afhænge af hvilke risici man ønsker at tage højde for.

PDF:

Mange undervarianter - Bagud konvertibelt - Under konstant udvikling

Mange kombinationsmuligheder - Formater omfattende og høj kompleksitet

Det er pt. ikke muligt at validere formatet

PDF/A:

Mange undervarianter - Bagud konvertibelt - Under konstant udvikling

Mange kombinationsmuligheder - Format er omfattende - Mindre komplekst end PDF

Kan delvist valideres



PDF/A archiving standard

<http://web.archive.org/web/20130502134821/http://www.adobe.com/enterprise/standards/pdfa/>

PDF/A is one set of standards among a suite of PDF-based standards managed by the International Organization for Standardization (ISO). It was developed to enable the long-term preservation of electronic documents and provides specifications for the creation, viewing, and printing of PDF documents, with the intent of preserving final documents of record as self-contained documents.

The standard does not define an archiving strategy or the goals of an archiving system. Rather, it identifies a "profile" for a PDF file that makes it **possible to reproduce the visual appearance of the document the exact same way in the future**. This profile specifies what must be included in the file, while prohibiting features that are not suitable for long-term archiving.



PDF/A archiving standard

Adoption by government and standards organizations

Since it became an ISO standard in 2005, PDF/A has gained significant acceptance. The earliest adopters have been government and standards organizations that define the regulatory standards businesses must live by. Here is a just a sampling of organizations that mandate, recommend, or accept PDF or PDF/A for the long-term archival of documents.



| | | |
|--|-------|-------------|
| U.S. National Archives and Records Administration (NARA) | PDF/A | Accepted |
| European Commission (MoReq) | PDF/A | Recommended |
| German government (SAGA v4) | PDF/A | Recommended |
| French government | PDF/A | Recommended |
| Dutch government | PDF/A | Mandated |
| National Archives of Sweden | PDF/A | Accepted |
| Austrian National Library | PDF/A | Recommended |
| National Archives of Norway | PDF/A | Recommended |
| Organization for the Promotion of Automated Accounting | PDF/A | Recommended |
| Brazilian federal legislature | PDF/A | Mandated |

PDF/A archiving standard

Adoption by government and standards organizations

Research studies and customer stories

The PDF file format is rapidly overtaking paper for long-term storage, as shown in these two market studies.

[AIIM](#) Market IQ study: [Content creation and delivery: The on-ramps and off-ramps of ECM](#)

90% of the organizations studied were using PDF for long-term storage of scanned documents.

89% were converting office files to PDF for distribution and archive.

In 5 years, use of paper for long-term storage drops to 77%, while PDF rises to 93%.

[PDF/A Competence Center](#) survey: [Nearly all archiving projects use PDF/A](#)

16% of 400 European respondents use PDF/A already.

50% intend to introduce it in the next 12 months.

Older archiving formats, such as TIFF, JPEG, and conventional PDF, decreased about 5% from the previous year's survey.



AFP/A archiving standard

http://afpcinc.org/afpa_release/



“This certification is a huge step for AFP and the AFP user community,” said Harry Lewis, President, AFPC. “Until the introduction of AFP/A, archival to an ISO-recognized format would require a transform to PDF.

While PDF/A is great for archiving PDF content, there is always the chance of introducing some small fidelity issue in the transform. If the resulting archive is incomplete or inaccurate, then integrity may be compromised, limiting the archive’s usefulness.

For the many users who run end-to-end AFP, the introduction of a fully realized, fully approved AFP/A format is a huge boon, providing increased efficiency and compliance with regulations that demand ISO-approved standards.”

PDF/A Check

PDF/A is based on the Portable Document Format technology developed by Adobe Systems Inc., but it removes those functions which are not suited to long-term archiving. These include audio and video elements and transparent objects.

An essential component for a PDF/A file's reproducibility is that the document is 100% self-contained.

These requirements ensure that the information needed to precisely reproduce the document every time is embedded within the file itself. This means that every PDF/A file contains all the content it needs, including text, raster images, vector graphics, fonts and color information.

<https://www.adobe.com> og Acrobat Pro

Ud over Adobes Acrobat kan du også finde PDF/A programmer her:

<https://www.luratech.com/en/products/pdfa-document-conversion-software-solutions/>

<https://www.foxitsoftware.com/downloads/>



PDF/A Check

<http://verapdf.org> (open source)



PDF/A validation

The specification for PDF/A is a set of restrictions and requirements applied to the “base” PDF standards (PDF 1.4 for PDF/A-1 and ISO 32000 for PDF/A-2 and PDF/A-3) plus a specific set of 3rd party standards.

veraPDF Implementation Checker

The Implementation Checker parses and analyzes PDF documents. It outputs two types of report: a report describing the PDF document and its metadata and a Validation Report describing conformance to PDF/A flavours.

AFP/A Check

<http://afpcinc.org/>

AFP/A validation

MO:DCA Reference (Mixed Object Document Content Architecture Reference)

AFRREBLK

The Implementation Checker parses and analyzes PDF documents. It outputs two types of report: a report describing the PDF document and its metadata and a Validation Report describing conformance to PDF/A flavours.

AFPC overvejer/tænker på et tool til validering.



Nordea 

??????????????

jan.hougaard.rasmussen@nordea.com

