

Vejledning i aflevering af forskningsdata bestående af dokumenter

Vejledning i hvordan I afleverer forskningsdata, som består af dokumenter i form af tekst, lyd, video eller billeder

Rigsarkivet juni 2022

Version 1.0

Indhold

0. Læsevejledning	2
A. Vejledningens målgruppe og anvendelse.....	2
B. Henvisning til øvrig vejledning.....	2
C. Lovgivning og retsfor skrifter.....	2
D. Definitioner.....	2
1. Format for aflevering	3
A. Indexfilerne afleveres i en mappe navngivet "Indices"	3
B. Kontekstdokumenter afleveres i en mappe navngivet "Contextdocumentation".....	3
C. Dokumenter konverteres til arkivformat og afleveres i en mappe navngivet "lbnr_Dokumenter" .	4
D. Oplysninger om digitale dokumenter registreres i et regneark	5

0. Læsevejledning

Offentlige myndigheder er forpligtet til at aflevere bevaringsværdige data og dokumenter til et offentligt arkiv.

Denne vejledning beskriver hvordan I afleverer forskningsdata, hvis der er tale om et studie, som består af dokumenter i form af tekst, lyd, video eller billeder. Det kan være et forskningsprojekt der har udført kvalitative interviews eller observationsundersøgelser.

A. Vejledningens målgruppe og anvendelse

Denne vejledning henvender sig til jer, der afleverer forskningsdata bestående af dokumenter.

Hvis der i afleveringen er flere dokumenter end I manuelt kan overkomme at registrere oplysninger om i et regneark, kan denne vejledning kun anvendes til skabelse af afleveringen, hvis I selv er i stand til at automatisere udfyldning af felterne i regnearket, herunder dokumenttitler og stier for placering af dokumenterne i mappestrukturen.

B. Henvisning til øvrig vejledning

Foruden denne vejledning har Rigsarkivet udarbejdet andre vejledninger, der har betydning for aflevering af dokumenter:

- Vejledning til bekendtgørelse om arkiveringsversioner nr. 128 (se 5 E-F samt H-J for krav til arkivformater)
- Vejledning om konvertering af dokumenter til TIFF

Følgende skabelon har betydning for aflevering af dokumenterne:

- Regneark til oplysninger om de digitale dokumenter (Skabelon)

I kan finde alt vejledningsmateriale på Rigsarkivets hjemmeside www.sa.dk.

C. Lovgivning og retsforskrifter

I kan finde information om lovgivning m.v. på Rigsarkivets hjemmeside www.sa.dk.

D. Definitioner

Afleveringsbestemmelse

Før afleveringen sættes i gang, udarbejder Rigsarkivet en bestemmelse, som fastsætter indholdet af afleveringen. Afleveringsbestemmelsen er en kravspecifikation for **indholdet** af afleveringen, mens bekendtgørelse om arkiveringsversionen fastsætter **formatet** for afleveringen, fx hvilke arkivformater som dokumenterne skal konverteres til inden aflevering.

Arkivformater Rigsarkivet benytter 6 arkivformater: TIFF, JPEG2000, MP3, WAV, MPEG2 og MPEG4.

Dokumenter i afleveringen er data, fx dagbøger, kvalitative interviews, transskriberede interviews, videooptagelser til feltstudier osv.

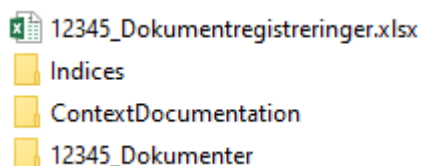
Kontekstdokumenter er dokumenter, der beskriver de afleverede data/dokumenter.

1. Format for aflevering

Et forskningsprojekt der består af dokumenter skal I aflevere på følgende måde (se figur 1):

1. Registreringer af oplysninger om alle digitale dokumenter placeret i mappestrukturen registreres i et regneark som navngives "lbnr_Dokumentregistreringer" og overholder denne vejlednings metadatastandard for dokumentsamlinger
2. Indexfilerne archiveIndex.xml og contextDocumentationIndex.xml afleveres i en mappe navngivet "Indices".
3. Kontekstdokumenter konverteres til arkivformat og afleveres i en mappe navngivet "Contextdocumentation"
4. Digitale dokumenter konverteres til arkivformat og afleveres som en kopi af en mappestruktur indeholdende dokumenterne, der placeres i en mappe navngivet "lbnr_Dokumenter"

Figur 1. Indhold på afleveringsmediet



Arkivet konverterer indexfilerne, kontekstdokumenterne, de digitale dokumenter og dokumentregistreringerne i regnearket til bevaringsformat, som er en arkiveringsversion jf. bestemmelserne i bekendtgørelse nr. 128 om arkiveringsversioner, bilag 1-8.

Afleveringsformatet specificeres nærmere herunder.

A. Indexfilerne afleveres i en mappe navngivet "Indices"

Rigsarkivet beder jer udfylde to spørgeskemaer med oplysninger om de dokumenter I skal aflevere og den medhørende dokumentation. Ud fra jeres oplysninger skaber vi to indexfiler i xml format, en arkivbeskrivelsesfil og en kontekstdokumentationsfil, som I skal have med i den endelige aflevering.

I vil modtage indexfilerne i en mail fra os, så I har dem, når I skal i gang med at lave afleveringen.

Når I skal skabe afleveringen placerer I indexfilerne i en mappe navngivet "Indices" på afleveringsmediet.

B. Kontekstdokumenter afleveres i en mappe navngivet "ContextDocumentation"

Rigsarkivet skelner mellem kontekstdokumenter og digitale dokumenter. Kontekstdokumenter betragtes som metadata, der beskriver de afleverede data, fx en projektbeskrivelse, en metoderapport, et spørgeskema eller variabelbeskrivelser. Andre dokumenter som kvalitative interviews, lydoptagelser af interviews eller videooptagelser til feltstudier betragtes derimod som data og skal afleveres som beskrevet under afsnit C.

I skal konvertere kontekstdokumenterne til arkivformat inden aflevering til arkiv. Følgende seks arkivformater er tilladte, jf. specifikationer, som fremgår af bekendtgørelse nr. 128 om arkiveringsversioner, 5 E-F samt H-J:

1. TIFF (.tif) til almindelige kontordokumenter i form af tekst, regneark og billeder
2. JPEG2000 (.jp2) til store tegninger, kort og fotos

3. MP3 (.mp3) til lydfiles
4. WAV (.wav) til lydfiles hvor høj kvalitet er vigtig
5. MPEG2 (.mpg) til videofiles
6. MPEG4 (.mpg) til videofiles

”Vejledning i konvertering til TIFF-format” med PDFCreator kan hentes fra Rigsarkivets hjemmeside www.sa.dk.

I placerer de konverterede kontekstdokumenter i en mappe navngivet ”ContextDocumentation” på afleveringsmediet.

Filnavne på kontekstdokumenterne skal stemme overens med dokumenttitlerne angivet i kontekstdokumentationsfilen (contextDocumentationIndex.xml, se afsnit A). Medfører dette for lange filstier kan I anvende en forkortet version af dokumenttitel eller dokumentID fra contextDokumentationIndex.xml filen til navngivning af filen.

Eksempel på dokumenttitel ”Afleveringsbestemmelse” i contextDokumentationIndex.xml filen:

```
<document>
  <documentID>1</documentID>
  <documentTitle>Afleveringsbestemmelse</documentTitle>
  <documentDescription>Afleveringsbestemmelse udarbejdet af
Rigsarkivet</documentDescription>
  <documentDate>2019-11-28</documentDate>
  <documentAuthor>
    <authorName>Jens Jensen</authorName>
    <authorInstitution>Rigsarkivet</authorInstitution>
  </documentAuthor>
  <documentCategory>
    <submissionInformation>
      <archivalProvisions>true</archivalProvisions>
    </submissionInformation>
  </documentCategory>
</document>
```

C. Dokumenter konverteres til arkivformat og afleveres i en mappe navngivet ”lbnr_Dokumenter”

I skal aflevere en kopi af den mappestruktur, der indeholder jeres digitale dokumenter i arkivformat. I placerer denne mappestruktur med dokumenter på afleveringsmediet i en mappe navngivet ”lbnr_Dokumenter”, hvor lbnr er løbenummeret for afleveringen udleveret af Rigsarkivet. Dette løbenummer fremgår af afleveringsbestemmelsen. Digitale dokumenter skal afleveres med originale filnavne.

I skal konvertere de digitale dokumenter til et af følgende seks arkivformater inden I afleverer dem til Rigsarkivet, jf. de specifikationer, som fremgår af bekendtgørelse nr. 128 om arkiveringsversioner, 5.E-5.F samt 5.H-5.J:

1. TIFF (.tif) til almindelige kontordokumenter i form af tekst, regneark og billeder
2. JPEG2000 (.jp2) til store tegninger, kort og fotos
3. MP3 (.mp3) til lydfiles
4. WAV (.wav) til lydfiles hvor høj kvalitet er vigtig

5. MPEG2 (.mpg) til videofiler
6. MPEG4 (.mpg) til videofiler

I kan placere de digitale dokumenter i jeres oprindelige mappestruktur, i en enkelt mappe eller i en nyoprettet mappestruktur, der eventuelt ordner og emneinddeler dokumenterne mere hensigtsmæssigt. Mappedstrukturens formål er at gruppere dokumenterne emnemæssigt, så de nemt kan fremsøges af en fremtidig arkivbruger.

”Vejledning i konvertering til TIFF-format” med PDFCreator kan hentes fra Rigsarkivets hjemmeside www.sa.dk.

D. Oplysninger om digitale dokumenter registreres i et regneark

I skal registrere oplysninger om de digitale dokumenter i et regneark i en fast skabelon. I finder skabelonen for dette regneark på Rigsarkivets hjemmeside www.sa.dk.

I skal navngive jeres udfyldte regneark ”lbnr_Dokumentregistreringer”, hvor lbnr er løbnummeret for afleveringen udleveret af Rigsarkivet, som fremgår af afleveringsbestemmelsen.

I afleverer regnearket med dokumentregistreringerne sammen med indexfilerne, kontekstdokumenterne og de digitale dokumenter på afleveringsmediet.

Regnearket indeholder kolonner til registrering af de 15 oplysninger fra metadatastandarden Dublin Core. Hertil er tilføjet oplysningen SourcePath. Metadatastandarden anvendt i skabelonen samt beskrivelser af hvordan oplysningerne udfyldes fremgår af figur 2.

I skal være opmærksomme på følgende, når I udfylder regnearket:

- I skal udfylde en række i regnearket pr. dokument i mappestrukturen.
- I skal bibeholde kolonner i regnearket i den angivne rækkefølge, også hvis I ikke udfylder dem.
- Kolonnenavne fremgår af række 1 i regnearket.
- Kolonnebeskrivelserne fra metadatastandarden i figur 2 fremgår af række 2 i regnearket.
- I må gerne tilføje egne eller redigerede kolonnebeskrivelser i række 3 i regnearket, så de mere præcist beskriver det faktiske indhold af kolonnerne.
- I skal udfylde kolonnerne **SourcePath** og **Titel** for alle dokumenter registreret i regnearket.
- I skal udfylde **Dato** for dokumentet, hvis den findes. Datoen skal overholde formatet ÅÅÅÅ-MM-DD, fx 2019-08-12.
- Øvrige oplysninger i regnearket er frivillige at udfylde. De registrerede oplysninger skal anvendes af fremtidige brugere af materialet til fremsøgning af et relevant dokument.
- I må gerne tilføje yderligere oplysninger om dokumenterne i ekstra kolonner indsat efter sidste kolonne i regnearket. I skal udfylde kolonnetitel i række 1 samt jeres egen kolonnebeskrivelse i række 3 for jeres ekstra tilføjede kolonner.

Figur 2. Vejledningens metadatastandard for dokumentsamlinger

*I skal udfylde oplysninger markeret med **

Kolonnenavn	Kolonnebeskrivelse
-------------	--------------------

SourcePath*	Angivelse af den fulde sti til dokumentet i mappestrukturen inkl. filnavn og ekstension, fx S:\ Skilsmisseprojekt\Børn og skilsmisse\Interview A – barn.tif
Title*	<i>(Titel)</i> Angivelse af original titel eller titel registreret efterfølgende på dokumentet. Dvs. billedets, filmens, videoens, lydoptagelsens eller tekstdokumentets titel.
Creator	<i>(Skaber)</i> Angivelse af navn på den der har taget initiativ til eller skabt dokumentets indhold. Fx navn på forfatter, interviewer, forsker, fotograf, forsøgsperson eller andre personer der har produceret materiale, der indgår som empiri i forskningen.
Subject	<i>(Emne)</i> Angivelse af emne eller kategori knyttet til dokumentet. Fx emneord der beskriver indholdet i dokumentet, navn på den mappe hvor dokumentet er placeret eller andre kategoriseringer, der grupperer dokumenterne.
Description	<i>(Indholdsbeskrivelse)</i> Fritekstbeskrivelse af dokumentets indhold. Fx hvad dokumentet indeholder, hvad filmen handler om eller hvad billedet viser.
Publisher	<i>(Udgiver)</i> Angivelse af udgiver af dokumentet. Fx filmselskab, producent, forlag. Sjældent relevant at udfylde for forskningsdata.
Contributor	<i>(Bidragyder)</i> Angivelse af navne på personer eller andre enheder, der har bidraget til skabelsen af indholdet af dokumentet. Fx medvirkende i en video, folk der taler på en lydoptagelse, interviewpersoner eller respondenter.
Date	<i>(Produktionsdato)</i> Angivelse af hvornår dokumentet er skabt. Fx oprettelsesdato for et dokument, dato for hvornår en video er optaget eller et interview er foretaget. Den angivne dato skal overholde formatet ÅÅÅÅ-MM-DD.
Type	<i>(Objekttype)</i> Angivelse af dokumentets type valgt fra følgende faste liste: Lyd, Video, Tekst, Billede.
Format	<i>(Originalformat)</i> Angivelse af dokumentets originale format. Her kan fx angives originalt medie, såsom tegning, fotografi, VHS-bånd eller kassettebånd. Her kan også angives det digitale format dokumentet havde før konvertering til bevaringsformat, fx mov, wma, excel regneark, word, pdf, docx eller jpeg.
Identifier	<i>(Identifikation)</i> Angivelse af et originalt ID for dokumentet anvendt i en given kontekst.
Source	<i>(Oprindelig kilde)</i> Angivelse af en reference til et andet materiale, som dokumentet er udledt af/digitaliseret fra. Fx angivelse af stregkode, navn, ID eller anden identifikation af materialet der refereres til.
Language	<i>(Sprog)</i> Angivelse af sprog anvendt i dokumentet.
Relation	<i>(Relation)</i> Angivelse af en reference til et relateret digitalt dokument/arkivalie i Rigsarkivets samling.

Coverage	<i>(Størrelse)</i> Angivelse af dokumentets mål. Fx størrelse på et billede hvor højde gange bredde er angivet i centimeter (cm) eller meter (m), et dokumentes størrelse angivet i Kilobytes (KB), Megabytes (MB) eller Gigabytes (GB) eller en videooptagelses spilletid angivet i minutter.
Rights	<i>(Rettigheder)</i> Angivelse af copyright ejer.