

# Quick-guide til produktion og test af en afleveringspakke

## Sådan kvalitetssikrer du statistikfilen før du anvender Hybris

Hybris udtrækker blot de data- og metadata der findes i statistikfilen.

Det er derfor vigtigt at du sikrer dig at kvaliteten af data- og metadata i statistikfilen lever op til Rigsarkivets krav, før du laver udtræk med Hybris.

Krav til data- og metadatakvalitet er listet herunder

- Tegnsættet for statistikfilerne skal være indkodet som UTF-8
- Der må ikke være linjeskift i indhold i sidste variabel
- Variabelnavne [Variable name] skal være angivet for alle variable.
- Variabelnavne skal være unikke inden for statistikfilen.
- Variabelnavne må aldrig begynde med et tal, men må godt være en blanding af bogstaver og tal
- Variabelnavne må max være 128 tegn lange
- Hvis variabelnavne er reserverede ord i SQL 99 (fx SELECT, END, DATE eller AGGREGATE) skal variabelnavnet omkranses af dobbeltapostrof
- Variabelbeskrivelser [Variable labels] skal være angivet for alle variable.
- Hvis der kun er angivet kodeforklaringer [value labels] for nogle af koderne [values] der fremgår af en skala, skal dette beskrives i variabelbeskrivelsen for variabelen.
- Der må konsekvent kun anvendes enten specialkoder (Special numeric) eller brugerdefinerede koder for manglende værdier.
- Specialkoder og brugerdefinerede koder for manglende værdier må kun anvendes for numeriske og kategoriske variable med datatyperne numerisk heltal og decimal.
- Hvis brugerdefinerede koder for manglende værdier er anvendt for variable med tekst, dato eller tidstempel skal dette beskrives i et kontekstdokument i afleveringspakken.
- Brugerdefinerede koder for manglende værdier skal altid fremgå af kodelisten [Value labels]
- Specialkoder for manglende værdier må kun antage værdier A-Z eller .a-.z.
- Datatyper for variable i statistikfilen skal overholde de tilladte datatyper i en afleveringspakke. Hvis dette ikke er tilfældet, skal du selv konvertere datatyperne til tilladte formater før udtræk laves med Hybris.
- Der skal være fuld overensstemmelse mellem datatype og længde for de variable, der indgår i en reference mellem flere statistikfiler i afleveringspakken.
- Alle koderværdier i data [Value] skal være beskrevet med en kodeforklaring [Value labels]
- Kodelister [Value labels] må kun være tilknyttet numeriske og kategoriske variable med datatyperne numerisk heltal, decimal og tekst.
- Koder [Values] skal være unikke inden for samme kodeliste [Value labels knyttet til en variabel]

Læs bilag 9 til bekendtgørelse 1007+ igennem, hvis du vil have en bedre ide om hvad en afleveringspakke er og hvilke krav der stilles til data- og metadatafilen

## Dette skal du have klar før du anvender Hybris

Du bør have nedenstående filer og informationer klar, før du anvender Hybris:

- Fremfinde de statistikfiler der skal afleveres til Rigsarkivet. OBS: Hvis personfølsomme oplysninger er adskilt fra statistikfilen i en nøglefil, skal både statistikfilen og nøglefilen indgå i afleveringspakken.
- Fremfinde løbenummer for afleveringspakken. Løbenummere er udleveret af Rigsarkivet og fremgår af afleveringsbestemmelsen.
- Udfylde indexfilerne archiveIndex.xml og contextDocumentationIndex.xml. Rigsarkivet har udviklet indtastningsværktøjer til udfyldning af indexfilerne. Se vejledningerne "Skab archiveIndex" og "Skab contextDocumentationIndex".
- Rigsarkivet skal godkende de udfyldte indexfiler, før du placerer dem i afleveringspakken.
- Fremfinde de kontekstdokumenter, der skal indgå i afleveringspakken. Disse fremgår af Rigsarkivets afleveringsbestemmelse.
- Konvertere kontekstdokumenterne til TIFF. Se hvordan i vejledningen om "TIFF-konvertering"
- Afklar om der findes unikke variable i statistikfilene som skal angives som nøglevariable under udtræk med Hybris
- Afklar om der skal dannes referencer mellem flere statistikfiler i afleveringspakken samt over hvilke variable disse referencer skal dannes.

## Disse dele af afleveringspakken kan du skabe med Hybris

- Navngive afleveringspakken med korrekt løbenummer.
- Danne en tom mappestruktur til afleveringspakken med struktur og navngivning af mapper, der overholder krav i bilag 9.
- Lave udtræk af data og metadata fra dine originale statistikfiler til data- og metadatafiler, der overholder de strukturelle krav til disse filer i bilag 9.
- Uploade archiveIndex.xml og contextDocumentationIndex.xml til afleveringspakken.

## Disse dele af afleveringspakken skal du selv lave

- Placere kontekstdokumenter i afleveringspakken i en rækkefølge der stemmer overens med rækkefølgen af dokumenter i contextDocumentationIndex.xml filen.

## Sådan tester du med testværktøjet Nemesis

- Klik på menupunktet NEMESIS og vælg den afleveringspakke, du vil teste. Klik på "Start".
- Læse fejlbeskeder fra testen i fejlloggen
- Nærstudere de krav i bilag 9 som fejlene omhandler, for at forstå hvordan fejlene rettes.

## Sådan retter du fejl i en eksisterende afleveringspakke

- Små fejl i metadatafilen kan rettes manuelt i en teksteditor
- Komplekse eller mange repeterende fejl i metadatafilen bør rettes direkte i statistikfilen, hvorefter der laves et nyt udtræk med Hybris.
- Fejl i datafilen rettes direkte i statistikfilen, hvorefter der laves et nyt udtræk med Hybris.
- Fejl i indexfilerne kan rettes i Rigsarkivets indtastningsværktøjer til disse. Værktøjerne validerer om filerne overholder krav til indexfiler. Ændringer i indexfilerne skal godkendes af Rigsarkivet.
- Fejl i kontekstdokumenter rettes ved at erstatte det eksisterende TIFF-dokument med et tilrettet.
- Tilføjes nye kontekstdokumenter til afleveringspakken som ikke fremgår af afleveringsbestemmelsen, skal afleveringsbestemmelsen opdateres med disse dokumenter og godkendes af Rigsarkivet.