

Depot N5 validering v1.0 IKAMR

torsdag 16. januar 2020 11.30

Navn	Depot N5 validering
Depot	IKAMR
Versjon	1.0
Dato	2020-02-14

Dette dokumentet er en fortsettelse og fordypning av "Depot mottak IKAMR".

- Type uttrekk: Noark 5

Del V: A4 Test og validering

- ⇒ All test og validering i Arbeidsområde på en kopi av mottatte originalfiler
- Liste over mulige testverktøy i avsnittene under
- Rapporter og loggfiler bevares sammen med skjermdumper
- Mal for test-logg og verktøy med logger er allerede kopiert i pkt. *A3-1 Data innhenting*
 - o `\repository_operations\analyse\..`
 - o `\repository_operations\..` (se de ulike filer vi har her)

A4-1 Oppretting av test-logg

- `..\repository_operations\FFKK_nnn_test-log.txt`
- Alle punkter i testfasen skal loggføres i test-logg med status:
 - o Godkjent
 - o Godkjent med anmerkning
 - o Avvises

A4-2 Teknisk konsistens, korrekt og komplett

A4-3 Struktur og formatkrav

A4-4 Teknisk- og bevarings-metadata

A4-5 Sjekksummer i enkeltfiler

*N5

- Noark 5 system
 - o Sjekksum til arkivuttrekk.xml (ADDML) skal følge med i uttrekket (typisk i info.txt)
 - o Sjekksummer på alle andre .xml og .xsd fra arkivuttrekk.xml
 - o Korrekte .xsd som publisert av Arkivverket
 - o Velformet .xml og validere .xml mot .xsd
 - o Innholdsanalyse (antall registreringer pr. år i arkivstruktur, offentlig- og løpende journal, endringslogg, dato- og datotid-felter, avskrivningsmåte)
 - o Generelle Noark 5-tester
 - o Spesifikke Noark 5-tester
 - o Validere arkivversjon av dokumenter
 - o Kryssreferanser
 - o Dokumentasjon
- `..\repository_operations\FFKK_nnn_Noark5-rapport.docx | .pdf (PDF/A)`

1. Innledende kopiering oppsett

a. Analyse

- Katalog `..\repository_operations\analyse\..` (kopiert under punkt a. over)
- Fast arbeidsdokument: `15KK_010_analyse.txt`
Original ligger i OneNote - IKAMR produksjonslinje, kanal arkiv-work: [arkiv-work N5 Innholdsanalyse v0.2](#)
Arbeidskopi som tekstfil ligger i mappestruktur (sjekk for oppdateringer jevnlig).

b. Arkade 5 info.xml fra Produsent

- Katalog `..\repository_operations\innhold\..` (kopiert under punkt a. over)
- Lastet ned fra orkidenett.no med zip-fil (valgfri): `15KK_nnn_info.xml_v1.5.1.0_orkidenett.xml`
- Pakket ut fra zip-fil sammen med tar-fil (valgfri): `15KK_nnn_info.xml_v1.5.1.0_zip.xml`
- Mottatt via epost (obligatorisk, journalføres): `15KK_nnn_info.xml_v1.5.1.0_epost.xml`

c. Arkade 5 rapport fra Produsent

- i. Katalog: `..\repository_operations\arkade5_v1.5.1.0_arkivskaper\..`
 - ii. Lastet ned fra orkidenett.no med zip-fil: [15KK_nnn_Arkade5_v1.5.1.0_arkivskaper_rapport.pdf](#)
 - iii. Kopiert fra `..\content\sip\content\administrative_metadata\repository_operations\..`
[15KK_nnn_Arkade5_v1.5.1.0_arkivskaper_rapport.xml](#)
[15KK_nnn_Arkade5_v1.5.1.0_arkivskaper_rapport.html](#)
- d. Noark 5 skjema
- i. Katalog: `..\content\schema\..` (kopiert under punkt a. over - ev. slette andre N5 skjemaer enn i dette uttrekket)
 - ii. Kun en Noark 5-versjonsmappe beholdes, resten fjernes når versjon er bekreftet
 - 1) Mappe `\arkivverket\..` inneholder xsd-skjema publisert fra Arkivverket
 - 2) xsd skjema levert med mottatt uttrekk
 f.eks `\ephorte-submission_v2.5.0\..` uttrekk eksempel under, tilsvarende for WebSak og Documaster
- a) Katalog for Noark 5 v3.1: `..\content\schema\noark5_v3.1\..`
 - i) `\arkivverket\..`
 - ii) `\ephorte-submission_v2.5.0\..`
 - iii) `\mod-1_ephorte\..`
 - b) Katalog for Noark 5 v4.0: `..\content\schema\noark5_v4.0\..`
 - i) `\arkivverket\..`
 - ii) `\ephorte-submission_v2.7.1\..`
 - iii) `\mod-1_ephorte\..`
 - c) Katalog for Noark 5 v5.0: `..\content\schema\noark5_v5.0\..`
 - i) `\arkivverket\..`
 - ii) `\documaster_v2.21.0\..`
 - iii) `\mod-1_documaster\..`

2. Arkade 5

- ⇒ <https://github.com/arkivverket/arkade5/releases>
- a. Katalog: `..\repository_operations\arkade5_v1.5.1.0\..`
 - b. Last inn uttrekk
 - c. Knapp "Start testing"
 - d. Skjermdump når test ferdig lagres som [15KK_nnn_Arkade5_v1.5.1.0.png](#)
 - e. Knapp "Vis rapport" åpner rapport i nettleser, lagres som [15KK_nnn_Arkade5_v1.5.1.0_rapport.pdf](#) med Adobe Acrobat
Pr. dato fungerer det ikke å lagre som PDF/A av ukjent årsak (test om MSWord utskrift til PDF/A fungerer?)
 - f. Kopier midlertidige rapportfiler fra Arkade 5 fra den valgte arbeidskatalogen som ble valgt for programmet, eks.:
`D:\uttrekk-arkade\arkade-tmp\work\<uuid>\administrative_metadata\repository_operations\..`
 arkade-log.xml lagres som [15KK_nnn_Arkade5_v1.5.1.0_rapport.xml](#)
 report.html lagres som [15KK_nnn_Arkade5_v1.5.1.0_rapport.html](#)
NB! Arkade sletter disse når program lukkes, så MÅ gjøres først!
 - g. Se over Arkade 5 testrapport og skrive notat i Analyse-dokument

3. KDRS XQuery

- ⇒ https://github.com/IKAMR/KDRS_Query/releases
- a. Katalog: `..\repository_operations\analyse\..`
 - b. Kjør siste oppdaterte IKAMR standardliste (ev. for gjeldende Noark 5 versjon og uttrekksleverandør)
 Original ligger i OneNote - IKAMR produksjonslinje, kanal arkiv-work: [xml-queries_noark5_2020-01-23](#)
 Arbeidskopi finnes i `..\repository_operations\analyse\xml-queries_noark5_2020-01-23.xml`
 - c. KDRS Query result (komplett fra standardliste b) over): [15KK_nnn_kdrs_query_results.txt](#)
 - d. IKAMR Xpath Runner (hardkodet data for kopiering til rapport): [15KK_nnn_ikamr_xpath_results.txt](#)
 - e. Generer valideringsrapport med data fra KDRS XQuery klar for videre redigering
 Mal finnes i `..\repository_operations\maler\IKAMR-Noark5-C-rapportmal_v1.2.0_2020-01-22.docx`
 Resultat dukker opp i `..\repository_operations\analyse\..`, men skal flyttes ut til `..\repository_operations\..`

4. oXygen XML Diff Directories

- a. Katalog: `..\repository_operations\analyse\..`
- b. Kopier uttrekks xsd skjema til (avhengig av Noark 5 versjon og leverandør/uttrekksprogram):
`..\content\schema\noark5_v3.1\ephorte-submission_v2.5.0\..`
`..\content\schema\noark5_v4.0\ephorte-submission_v2.7.1\..`
`..\content\schema\noark5_v5.0\documaster_v2.21.0\..`
- c. Kjør Diff Directories mot (avhengig av Noark 5 versjon og leverandør/uttrekksprogram):
`..\content\schema\noark5_v3.1\arkivverket\..`
`..\content\schema\noark5_v4.0\arkivverket\..`
`..\content\schema\noark5_v5.0\arkivverket\..`
- d. Dokumenter avvik i analyse arbeidsdokument iht. maltekster der for xsd skjema Diff Directories

- e. Vurdere å ta skjermdump av viktige avvik (vurderes mot teksten i analyse tekstfila)
5. oXygen XML Editor XQueries
- Dokumenter avvik i analyse arbeidsdokument iht. maltekster der for de kjente problemområder
- z. Utgangspunkt
 - i. Katalog: `..\repository_operations\analyse\..`
 - ii. Standardspøringer
Original ligger i OneNote - IKAMR produksjonslinje, kanal arkiv-work: [arkiv-work N5 Innholdsanalyse v0.2](#)
Arbeidskopi som tekstfil ligger i mappestruktur (sjekk for oppdateringer jevnlig).
 - a. arkivstruktur.xml
 - i. XQuery for antall oppføringer pr. år
 - ii. XQuery for avskrivningsmaate
 - b. Endringslogg.xml
 - i. XQuery for antall oppføringer pr. år
 - c. loependeJournal.xml
 - i. XQuery for antall oppføringer pr. år
 - d. offentligJournal.xml
 - i. XQuery for antall oppføringer pr. år
6. Documaster Noark Extractor Validator
- ⇒ <https://github.com/KDRS-SA/noark-extraction-validator/releases>
- a. Katalog: `..\repository_operations\noark-extractor-validator\..`
 - b. Vurdere behov for `..\mod-1_ephorte\..`. Ut i fra oXygen XML Editor XQueries over
 - i. F. eks. metdatakatalog.xsd hvis der er verdier av avskrivningsmaate som ikke er blant de 7 gyldige
 - ii. F. eks. arkivstruktur.xsd hvis fortsatt Arkade5 og Noark Extractor Validator bruker feil (konverteringsverktøy)
 - c. Startes fra `C:\prog\noark-extractor-validator-0.7.0_2020-01-14\..`
alternativt ligge på disk med uttrekket eks. `I:\prog\..`
 - d. Java kompilert file har navn `noark-extraction-validator-0.7.0-IKAMR_2020-01-14.jar`
`Y:\elarkiv\software\documaster\ikamr\noark-extraction-validator-0.7.0-IKAMR_2020-01-14\noark-extraction-validator-0.7_IKAMR_2020-01-14.jar`
 - e. Kjøres fra .bat fil med navn
`15KK_nnn_n5-validate-IKAMR_2020-01-14_I.bat` <<< stasjonsnavn_I for katalog med uttrekk `I:\15KK_nnn_n5-validate-IKAMR_2020-01-14_I_schema.bat` <<< tar med _schema hvis CUSTOM brukt
Merk at mal/tidligere .bat fil settes 'REM exit' før kjøring, og tilbake til 'exit' når ferdig
 - f. Kjøres til Excel (ikke med XML no fordi den får feil, ikke komplett og feil tegnsat og krasjer)
 - g. Aktiver 'exit' øverst i .bat fil og kopier denne .bat fil til dette punktets hovedkatalog
 - h. Åpne Excel regneark, redigere første fane til gode marger og ev. ekstra versjonsinfo program
 - i. Sjekk avvik hovedpunkter (opp mot avvik oXygen XML Editor XQueries over og for øvrig)
 - j. Dokumenteres i analyse tekstdokument
7. DROID
- ⇒ <http://digital-preservation.github.io/droid/>
- ⇒ <https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>
- a. Katalog: `..\repository_operations\droid_v6.4\..`
Sjekk jevnlig for nyere versjoner av DROID, samt Binary and Container Signature files pr. 03.02.2020:
 - i. DROID v6.4
 - ii. Binary Signature File: DROID_SignatureFile_V96
 - iii. Container Signature File: container-signature-20200121
 - b. Lagre filer
 - i. `15KK_nnn_droid_v6.4_V96.droid`
 - ii. `15KK_nnn_droid_v6.4_V96.csv`
 - iii. `15KK_nnn_ephorte-SAK3_report-planets.xml`
 - iv. `15KK_nnn_ephorte-SAK3_report-droid.xml`
 - v. `15KK_nnn_ephorte-SAK3_report-droid.pdf`
 - vi. Dokumentere i analyse tekstdokument
8. veraPDF
- ⇒ <https://verapdf.org/software/>
- a. Katalog: `..\repository_operations\verapdf_v1.14.2\..`
 - b. Alternativ hvis v1.14.2 krasjer, Katalog: `..\repository_operations\verapdf_v1.12.1\..na`
 - c. Lagre filer
 - i. `15KK_nnn_verapdf_v1.14.2_htmlReport.html`

- ii. [15KK_nnn_verapdf_v1.14.2_xmlReport.xml](#)
- iii. [15KK_nn_verapdf_v1.14.2.png](#)
- iv. Åpne html-fil i nettleser og hvis mulig skriv ut til PDF
[15KK_nnn_verapdf_v1.14.2_report.pdf](#)
- v. Dokumentere i analyse tekstdokument

9. KOST-Val

- ⇒ https://kost-ceco.ch/cms/index.php?kost_val_de
- ⇒ <https://github.com/KOST-CECO/KOST-Val/releases>
- a. Katalog: `..\repository_operations\kost-val_v1.9.6\..`
- b. Alternativ hvis v1.9.6 krasjer, Katalog: `..\repository_operations\verapdf_v1.8.4\..`
- c. Lagre filer (kopieres fra temp-katalog til KOST-Val, lage link til denne mappe på skrivebordet)
 - i. [15KK_nnn_kost-val_v1.9.6_log.xml](#)
 - ii. [15KK_nnn_kost-val_v1.9.6_a.png](#) ta skjermdump ved start kjøring hoveddialog
 - iii. [15KK_nnn_kost-val_v1.9.6_b.png](#) ta skjermdump dialog for ferdig kjøring, som oftest med avvik
 - iv. Åpne xml-fil i nettleser og hvis mulig skriv ut til PDF (NB! Store XML-filer åpner seg kanskje aldri...)
[15KK_nnn_kost-val_v1.9.6_log.pdf](#)
 - v. [kost-val.xls](#) (med i standardkatalog kopiert over)
- vi. Dokumentere i analyse tekstdokument

10. KDRS Analyse

- ⇒ https://github.com/IKAMR/KDRS_Analyse << ikke noen offisiell release ute enda
- a. Katalog: `..\repository_operations\analyse\..`
- b. Last inn: info.xml
- c. Last inn: DROID
- d. Last inn: veraPDF
- e. Last inn: KOST-VAL
- f. (Last inn: Arkade 5.xml < ikke ferdig?)
- g. (Last inn: Documaster Noark Extractor.xml < ikke ferdig? + problemer xml generering)
- h. [15KK_nnn_kdrs-analyse_v.9.6_log.xml](#)
- i. Dokumentere i analyse tekstdokument

11. Noark 5 Word rapport

- a. Mal: IKAMR-Noark5-C-rapportmal_v1.1.2_2018-11-30.docx
- b. Rapport (opprettes av KDRS XQuery):
`..\repository_operations\15KK_nnn_Noark5_rapport.docx`

12. Noark 5 Excel logg

- a. Mal: IKAMR-Noark5-D-validator-logg_v1.0.1_2018-06-21.xlsx
- b. Logg (ønskes automatisert):
`..\repository_operations\15KK_nnn_Noark5_validator-log.xlsx`

13. Elarkiv sjekkliste

- a. Mal: IKAMR-21_Sjekkliste_earkiv_godkjenning_v2.2.0_2020-01-10.docx <<< Ny versjon med tabell anker
- b. Rapport (ønskes automatisert):
`..\repository_operations\15KK_nnn_Elarkiv_sjekkliste.xlsx`

Del VI: Opprette AIP

A5; 2018-mm-ddThh:nn; Oppretting AIP; ;

Del VII: Lagre AIP

A6; 2018-mm-ddThh:nn; Lagre AIP; ;