



RIGSARKIVET

Ordforklaringer

Rigsarkivets politik og strategi for digital arkivering



ORDFORKLARINGER

Rigsarkivets politik for digital arkivering

Rigsarkivets strategi for digital bevaring 2025

Afleveringspakke: (SIP) En informationspakke, der er afleveret til Rigsarkivet af dataskaber. Pakkens struktur og indhold skal overholde gældende regler for en afleveringspakke for den pågældende indholdstype.

Anvendelsespakke: (DIP) En informationspakke afledt fra en eller flere bevaringspakker og gjort tilgængelig af Rigsarkivet efter anmodning fra en bruger. En anvendelsespakke kan indeholde data og dokumenter, som er konverteret til mere brugbare anvendelsesformater end bevaringsformaterne i selve bevaringspakkerne.

Anvendelsesformat: Et format udleveret til brugeren, der er mere anvendeligt end selve bevaringsformatet. Eksempelvis bevares statistiske data som en relational database i SIARD format i bevaringspakken, men disse data kan konverteres til et statistikformat ved udlevering til en bruger og dermed lettere anvendes i gængse statistikprogrammer.

Arkiveringsprocessen: En proces, der dækker aktiviteter i alle dele af datas livscyklus: Fra skabelse af data over aflevering til arkiv, bevaring, tilgængeliggørelse og til genanvendelse.

Bevaringspakke: (AIP) En informationspakke, som er overført til bevaringslageret i et OAIS-arkiv. Når en afleveringspakke er valideret, godkendt og placeret på bevaringslageret af Rigsarkivet skifter den status til en bevaringspakke.

Bevaringsformat: Et format som Rigsarkivet har godkendt til langtidsbevaring.

Bevaringsplan: En plan, der er udarbejdet af Rigsarkivet for bevaring af en indholdstype. Bevaringsplanen indeholder fx et eller flere godkendte bevaringsformater for indholdstypen, angivelse af bevaringsniveau samt specificering af valideringskrav.

Bevaringsniveau: Bevaringsniveauer er forskellige modeller for bevaring af data. Rigsarkivets bevaringsniveauer er endnu ikke færdigdefineret, men nedenstående eksempel på bevaringsniveauer illustrerer forskellige typer af bevaringsniveauer.

1. Rigsarkivet modtager data i godkendt bevaringsformat med vedligeholdelse.
2. Rigsarkivet modtager data både i originalformat og bevaringsformat. Sidstnævnte med vedligeholdelse.
3. Rigsarkivet modtager udelukkende originalkopi uden vedligeholdelse (men evt. overvågning).
4. Rigsarkivet modtager ikke data. Data bevares evt. i institution

Bevaringssikkerhed: Bevaringssikkerhed kan både betyde at data lagres og håndteres sikkert så datas fortrolighed ikke kompromitteres. Det kan også betyde at aktive bevaringsindsatser sikrer at data kan tilgås og forstås i fremtiden. Anvendelse af forskellige bevaringsniveauer medfører varierende bevaringssikkerhed.

Data: Se information.

Dataskaber: Den aktør, f.eks. myndighed eller privat virksomhed eller forening, som skaber data, der afleveres til og bevares af Rigsarkivet. Også kaldet "arkivskaber".

Digital bevaring: Digital bevaring defineres som en aktiv og bevidst håndtering af digitale data gennem hele deres livscyklus, der sikrer at de kan tilgås og anvendes i fremtiden.

Digitalt skabte data: Data, som er skabt digitalt. Det kan være data registreret i et offentligt it-system eller data indsamlet eller produceret i digital form til

et forskningsprojekt. Det er også dokumenter oprindeligt skabt i digitale formater såsom Word og PDF eller digitale dokumenter skabt i forbindelse med et skanningsprojekt.

FAIR: FAIR står for Findable, Accessible, Interoperable og Reuseable. Principperne beskriver de vigtigste egenskaber, som data, metadata og elementer i en infrastruktur skal have for at data og metadata kan findes, genanvendes og udveksles på tværs.

Genanvendelse: De handlinger en arkivbruger udfører, når brugeren opnår adgang til og anvender data. Fx fortolkning af data eller dokumenter eller analyse af data i et datasæt. Følgende forudgående aktiviteter i arkiveringsprocessen muliggør at data kan genanvendes:

- Når data er **disponible** (available) kan de tilgås.
- Tilvejebringelse af metoder og værktøjer sikrer at data kan **læses** (retrievable).
- Hvis data er disponible og kan læses er informationen **tilgængelig** (accessible).
- Beskrivelse af datas kontekst og metadata sikrer, at informationen er **forståelig** (understandable) og dermed **genanvendelig** (usable).

Indholdstype: En indholdstype er en gruppe af filformater, der har til formål at udføre de samme funktioner. F.eks. er video en indholdstype med mange forskellige videofilformater under sig.

Information: Objektet, som Rigsarkivet bevarer, er først og fremmest information. Der skelnes mellem data, information og viden på følgende måde:

- **Data:** Data udgør de tegn, filer og bitstrømme som håndteres i Rigsarkivet.
- **Information:** Data kan repræsenteres på forskel-

lige måder, fx. i forskellige strukturer, formater og datatyper, men stadig indeholde den samme information. Datas informationsindhold skal ofte forstås i lyset af den metadata og kontekstdokumentation, der er knyttet til data. Data kan således indeholde samme semantiske betydning, selvom repræsentationen af data er forskellig. Migrering af data kan ændre datas repræsentation.

- **Viden:** Når en bruger fortolker på informationsindholdet i data skabes ny viden.

Informationspakke: En informationspakke er den struktur, ofte en mappestruktur, de arkiverede data, metadata og kontekstdokumenter placeres og bevares i. Informationspakken kan indeholde obligatoriske filer med pakkeinformation, som beskriver hvor det forskellige indhold i pakken er placeret.

Integritetstjek: En handling, der beregner en unik værdi (checksum) for data og validerer checksummen imod den oprindelige værdi for data. Hvis den nye beregnede checksum er forskellig fra den oprindelige checksum er data angiveligt ændret.

It-system: Et it-system bruges til at skabe og håndtere data og dokumenter på en struktureret og sikker måde. Data afleveret til Rigsarkivet er som oftest skabt eller lagret i et it-system.

Konvertering: Overførsel af data fra et filformat til et andet filformat. Konvertering af data er en naturlig del af migreringsstrategien.

Lagringsstruktur: Informationspakkens mappehierarki og placeringen af elementer og relationen mellem disse i informationspakken.

Metadata: Metadata er den information, der beskriver data og giver dem betydning. Fx 1955 er en række cifre uden betydning. Hvis metadata beskriver at værdien 1955 har betydningen ”Årstal for opførelse af bygning” er denne værdi blevet tildelt en specifik betydning. Der findes forskellige typer af metadata for forskellige niveauer, fx beskrivende metadata anvendt til søgning i søgekataloget til Rigsarkivets samling eller datanær metadata der beskriver indhold i tabeller og kolonner i en informationspakke.

Migrering: En handling, der flytter data fra en lagringsstruktur til en ny.

OAIS: Open Archival Information System er en international standard til beskrivelse, af hvad et digitalt arkiv er, og hvordan arkivet fungerer.

Tilgængeliggørelse: Den aktivitet i arkiveringsprocessen hvor data fra bevaringspakken konverteres til et anvendelsesformat som udleveres til brugeren i en anvendelsespakke eller tilgås direkte genrejst i en egnet klient, som brugeren har adgang til. Klienten kan være et program, som Rigsarkivet stiller til rådighed eller gængs software brugeren selv har installeret på sin computer. Arkivbrugere genanvender data i klienten.

Troværdighed: De data Rigsarkivet bevarer skal indeholde troværdig information. Troværdighed dækker at informationen er pålidelig, autentisk og integriteten intakt.

- At information er **pålidelig** betyder, at man kan stole på, at informationen er en fuldstændig og nøjagtig gengivelse af de transaktioner, aktiviteter eller fakta som informationen vidner om.
- At information er **autentisk** betyder, at det er bevist at informationen er det som den foregiver

at være, skabt eller sendt af den person, som er angivet som afsender af informationen på det angivne tidspunkt.

- At informationens **integritet** er intakt betyder, at informationen er komplet og uændret.

Da migreringsstrategien medfører risiko for datatab under migrering er sikring og dokumentation af datas troværdighed central for Rigsarkivet.

Udtræksmodeller: Forskellige modeller for hvordan data kan udtrækkes til en afleveringspakke. Modeller kan fx være at afleveringspakken indeholder:

- et udvalg af tabeller fra ét system.
- et udtræk af de væsentligste informationer, som samles i nye tabeller.
- en samling af data fra flere systemer
- en opsplitning af data fra samme system i flere afleveringspakker
- et udtræk til en på forhånd defineret udtræksstandard for systemtypen

Validering: En handling, der bekræfter nøjagtigheden af data eller sikrer at indholdet af et filformat stemmer overens med specifikationen for dette filformat.

Vigtige egenskaber: Vigtige egenskaber (significant properties) er egenskaber ved et format, som er vigtige at bevare for eftertiden, så data kan anvendes til det ønskede formål. F.eks. er det vigtigt at bevare information om formler anvendt i et regneark, hvis en bruger i fremtiden skal kunne kontrollere, om beregningerne i regnearket er udført korrekt.



RIGSARKIVET

www.sa.dk