



RIGSARKIVET

Strategi for digital bevaring 2025



INDHOLD

Indledning	3
Strategi for digital bevaring	4
Indsatsområder i arkiveringsprocessen	5
Data skabes - indsamling	5
Data bevares - modtagelse og langtidsbevaring	5
Data genanvendes - tilgængeliggørelse	6
Tværgående indsatsområder	7
Bevaringsplanlægning og -metodik	7
Samarbejde og standardisering	7
Troværdige data	8
Innovation	8

INDLEDNING

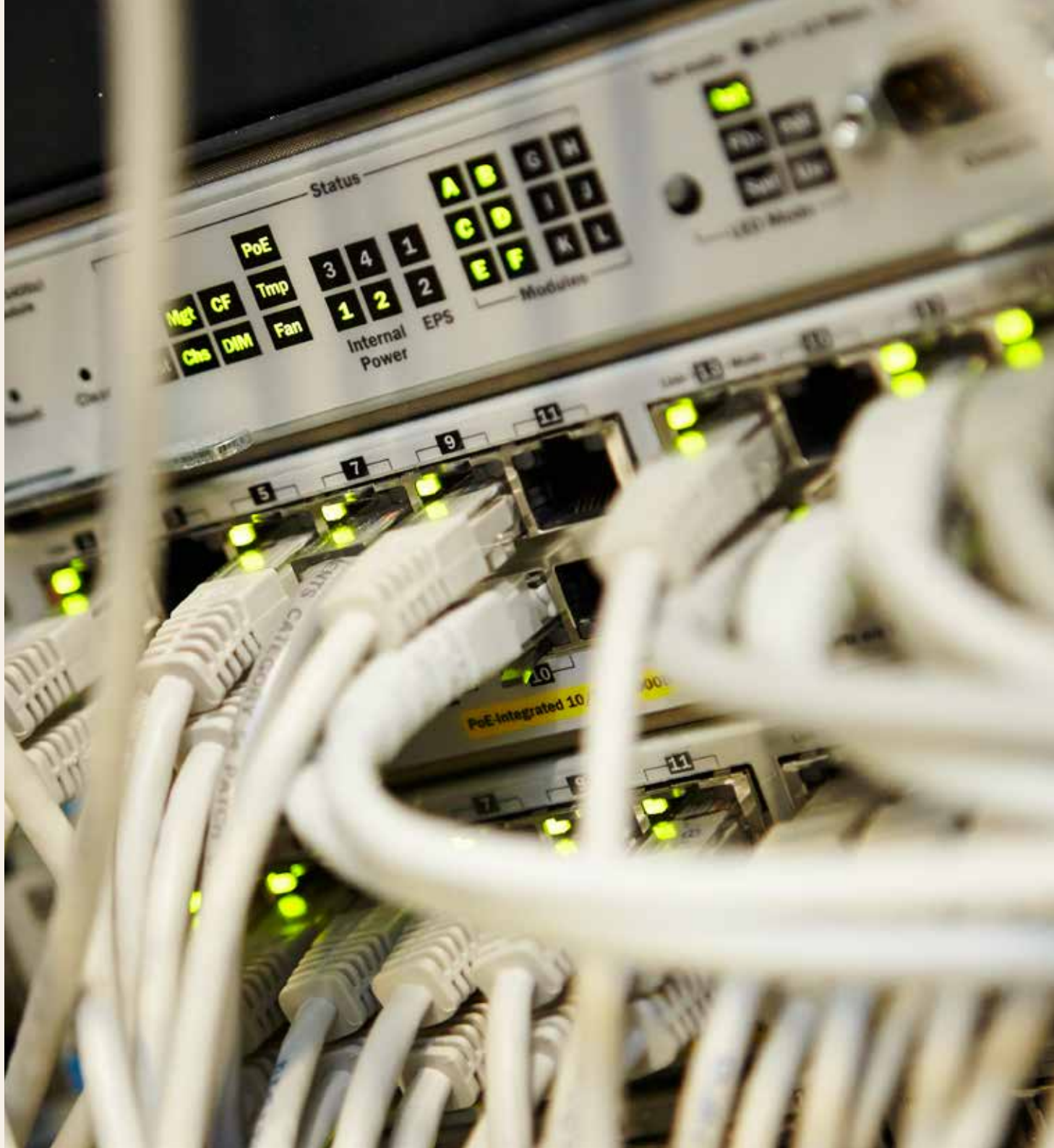
Dette dokument beskriver Rigsarkivets strategi for digital bevaring. Strategien præsenterer ud fra rammerne og pejlemærkerne i ”Rigsarkivets politik for digital arkivering” vores prioriterede indsatsområder og forslag til, hvor vi konkret vil implementere disse i de kommende år.

Denne strategi erstatter ikke Rigsarkivets migreringsstrategi, men supplerer den. Vi vil fortsat basere os på migreringsstrategien med anvendelse af få langtidsholdbare bevaringsformater, der sikrer, at data kan tilgås og genanvendes også på langt sigt.

Vi ønsker at supplere migreringsstrategien med nye tiltag, der opvejer nogle af migreringsstrategiens ulemper, fx at der kan ske bevidste eller ubevidste tab af data eller formatters vigtige egenskaber under migrering. De nye tiltag har til formål at sikre, at vi kan bevare mere end tidligere og optimere genanvendeligheden.

Rigsarkivets vision

”Rigsarkivets arbejde med digital arkivering skal sikre en omkostningseffektiv bevaring af flere typer af bevaringsværdige data på en måde, der sikrer større genanvendelighed for brugerne og muliggør en hurtigere adgang til data”.



STRATEGI FOR DIGITAL BEVARING

Rigsarkivets strategi for digital bevaring indeholder fire strategiske mål med tilhørende indsatsområder.

Digital bevaring omhandler blandt andet at fastsætte regler for bevaringsformater og informationspakkers struktur og indhold, og at etablere en sikker infrastruktur til at modtage, lagre og tilgængeliggøre digitalt skabte data.

Digital bevaring hænger samtidig tæt sammen med øvrige aktiviteter i digital arkivering, der dækker alle dele af datas livscyklus: Fra data bliver skabt og derefter afleveret til arkiv, hvor data bevares og tilgængeliggøres, og til de senere genanvendes, når der i samfundet opstår et behov for at anvende data igen.

Nogle af bevaringsstrategiens indsatsområder kan have betydning for andre faser i digital arkivering end blot i bevaringen af data. Derfor uddybes de nye indsatsområder nedenfor dér, hvor de får betydning i og på tværs af digital arkivering.

Prioriteringen af strategiens indsatsområder tager udgangspunkt i pejlemærkerne fra ”Rigsarkivets politik for digital arkivering”:

1. Genanvendelighed
2. Troværdighed
3. Sikkerhed
4. Flexibilitet
5. Innovation
6. Samarbejde
7. Standardisering
8. Transparens

Vi sikrer dokumentationen af det danske samfund for både nutiden og fremtiden

Vi sætter data i spil, så de skaber mere værdi for flere

Vi bidrager til effektivisering af den offentlige sektor

Strategi for digital bevaring 2025

Strategiske mål			
Hurtigere og bedre genanvendelse af data	Fælles vedligehold af viden og værktøjer	Bevaring af flere typer af bevaringsværdige data	Effektiv og sikker håndtering af voksende datamængder
Indsatsområder			
Udstede nye bekendtgørelser, der understøtter, at data er FAIR	Anvende fælles europæiske bevaringsformater	Anvende ny metode til fleksibel tilføjelse af bevaringsformater	Implementere nyt sikkert system til modtagelse og bevaring af data
Skabe grundlag for genanvendelse af originale data	Samarbejde om vidensopbygning, metoder og værktøjsudvikling	Anvende flere bevaringsniveauer	Anvende og udvikle innovative teknologier
Dokumentere processer udført på data i hele deres livscyklus	Inddrage brugere i vurderingen af formaters vigtige egenskaber	Indføre mulighed for bevaring af både original og bevaringsformat	Analysere mulighed for løbende aflevering
Udstille data og metadata i værktøjer, der optimerer søgning i og på tværs af data	Samarbejde om standarder og services i den fællesoffentlige infrastruktur	Indføre udtræksmodeller til fleksibelt design af informationspakke	Analysere mulighed for sikker digital aflevering

Figur 1: Rigsarkivets strategi for digital bevaring

INDSATSOMRÅDER I ARKIVERINGSPROCESSEN

Data skabes - indsamling

Bevaringsværdige data skal så vidt muligt identificeres på et tidligt tidspunkt. Det giver dataskaberne mulighed for at indrette it-systemer med bevaringsværdige data sådan, at data kan afleveres til arkiv efter gældende regler.

Statslige myndigheders bevaringsværdige data identificeres primært ved, at myndighederne anmelder nye it-systemer til Rigsarkivet. Hvis Rigsarkivet vurderer, at data skal bevares, gennemgår systemet en godkendelsesproces. Denne proces sandsynliggør, at aflevering af data til arkiv ikke vil give anledning til problemer. Processen sikrer også en tidlig dialog med dataskaberne om eventuelle udfordringer i forhold til aflevering af data.

For statslige myndigheder, herunder sektorforskningsinstitutioner og universiteterne er der udstedt en bekendtgørelse for anmeldelse af forskningsdata under bestemte forudsætninger. For kommunale og regionale myndigheder har Rigsarkivet udstedt generelle bevaringsbestemmelser, som dataskaberne selv skal vurdere nye it-systemer ud fra.

Anvende ny metode til fleksibel tilføjelse af bevaringsformater

Når det identificeres, at it-systemer indeholder bevaringsværdige data, som ikke umiddelbart kan afleveres til arkiv i overensstemmelse med gældende regler, vil vi implementere en ny metode til systematisk at vurdere nye filformater til langtidsbevaring.

I arbejdet med at forberede aflevering af data bestemmer Rigsarkivet informationspakkens tidsmæssige

afgrænsning, model for aflevering og indhold af data og dokumentation. Den informationspakke, som dataskaberne afleverer til arkiv, skal indeholde tilstrækkelig dokumentation af data og deres kontekst til, at de kan anvendes og fortolkes på både kort og langt sigt.

Rigsarkivet skal i den forbindelse strukturere dokumentere beslutninger om udvælgelse af data og konvertering samt datas ophavssituation, og sikre, at disse oplysninger er transparente for brugerne.

Indføre udtræksmodeller til fleksibelt design af informationspakke

Udvælgelse af data betyder, at indholdet af en informationspakke kan være en 1:1-version af it-systemets data, men det kan også være et udvalg af databasens tabeller, et udtræk af de væsentligste data eller en sammenstilling af data fra flere forskellige it-systemer. Rigsarkivet vil arbejde bevidst med at definere og dokumentere de forskellige typer af udtræksmodeller, så informationspakkens tilblivelse bliver så transparent som muligt for brugerne.

Det er vores målsætning, at dokumentation af informationspakker – specielt i form af strukturerede metadata – så vidt muligt lever op til internationale standarder ift. beskrivelse af informationspakkens struktur og indhold.

Samarbejde om standarder og services i den fællesoffentlige infrastruktur

Rigsarkivet vil i strategiens levetid undersøge mulighederne for ”arkivering gennem design”. Det indebærer blandt andet, at dataskabernes it-systemer med bevaringsværdige data indrettes på en sådan måde, at muligheden for at eksportere data i standardformater til offentligt arkiv er tænkt med fra starten. Samtidig vil vi have et skarpere fokus på at indgå i og bidrage til den fællesoffentlige infrastruktur ved at anvende nationale standarder som fx DCAT-AP-DK til beskrivelse af datakataloger, -sæt og -tjenester og den fællesoffentlige standard for beskrivelse af it-systemer samt tilbyde adgang til data gennem borger.dk.

Data bevares - modtagelse

Rigsarkivet modtager data fra hundredevis af dataskabere hvert år. Det er fx statslige myndigheder, forskningsinstitutioner og en lang række kommunale, regionale og private dataskabere. Data skabes i mange forskellige formater, strukturer og datamodeller. Afleveringen af data til Rigsarkivet sikrer, at data bevares efter en norm, der giver bevaringssikkerhed, dvs. sikrer datas tilgængelighed over tid. Data kan repræsenteres på forskellige måder. Det kan fx have forskellige strukturer og formater, men stadig indeholde den samme information. Rigsarkivet bevarer derfor først og fremmest information. Data bevares hovedsageligt i en anden form end den originale, fordi data skal migreres til informationspakker med særlig struktur, formater og dokumentation, der skaber omkostningseffektiv bevaring under hensyntagen, til at teknologier forældes.

Rigsarkivet validerer dataskabernes informationspakker, når de modtages. Denne test af, hvorvidt data er afleveret efter Rigsarkivets specifikationer, giver størst mulig sikkerhed for, at data er afleveret korrekt og forbliver tilgængelige i fremtiden. Rigsarkivet overvåger løbende de afleverede data og migrerer dem, når det vurderes, at data er i fare for teknologisk forældelse, eller hvis Rigsarkivet beslutter at anvende andre standarder i sin bevaringsmetodik.

Når Rigsarkivet har kontrolleret afleveringspakken, bliver den overført til langtidsbevaring i Rigsarkivets samlinger.

For at imødegå trusler mod datas integritet og tilgængelighed baserer Rigsarkivet sin fysiske bevaring af data på distribueret digital bevaring. Det betyder, at data findes i flere eksemplarer på flere medieteknologier og på flere lokaliteter og med en organisatorisk adskillelse imellem eksemplarerne. Langtidsbevaringen inkluderer også regelmæssige tjek af datas integritet, så Rigsarkivet er sikker på, at data ikke ændrer sig imellem migreringerne.

Implementere nyt sikkert system til modtagelse og bevaring af data

Rigsarkivet vil investere i et nyt system til at modtage og bevare data. Systemet skal sikre skalérbarhed, reducere personafhængigheder, anvende standardiseret software og forenkle og automatisere processer.

Data genanvendes - tilgængeliggørelse

Der er et stort potentiale for genanvendelsen af bevaringsværdige data, som er afleveret til arkiv i overensstemmelse med Rigsarkivets standarder for veldokumenteret, systemuafhængig arkivering.

Det er en målsætning at sikre, at arkiverede data bliver så lette at genanvende som muligt. Det skal ske ved, at vores fremtidige metodeudvikling og regelfastsættelse bygger på arkivets og brugernes konkrete erfaringer med at tilgængeliggøre og genanvende arkiverede data.

Udstede nye bekendtgørelser, der understøtter, at data er FAIR

Rigsarkivet arbejder løbende for at sikre tilstrækkelige metadata, så udstilling af metadata og data følger FAIR-principperne. Det betyder, at Rigsarkivet tilslutter sig målsætningen om, at data skal være Findable, Accessible, Interoperable og Reuseable, og at dette skal sikres bedre i alle led i datas livscyklus - både når data skabes, afleveres, bevares og genanvendes.

Rigsarkivets behandling af bevaringspakken resulterer i en anvendelsespakke, der stilles til rådighed for en bruger ud fra et ønske om at få så mange data ud så hurtigt som muligt. Det gælder:

- Hurtig og automatisk udstilling af frit tilgængelige data på nettet
- Anonymisering af informationspakker
- Udstilling af data på nye måder
- Digital udlevering af data til brugere

Udstille data og metadata i værktøjer, der optimerer søgning i og på tværs af data

Med velovervejede indsatser igennem hele processen for digital arkivering vil vi sikre, at anvendelsespakken skabes og kan anvendes i overensstemmelse med brugernes behov. Rigsarkivet har fokus på, at data skal være søgbare, og at data fra forskellige kilder kan kombineres og bidrage til at skabe ny viden ved at understøtte interoperabiliteten mellem informationspakker.

Vi vil også, så tidligt som muligt i processen, forsøge at gennemføre indsatser, som letter anonymisering af data. Det gør vi for at sikre, at data lettere og hurtigere kan blive tilgængelige. Adgang til data forudsætter dog altid, at data er tilgængelig jf. arkivloven, og at data dermed ikke havner uvedkommendes hænder.

TVÆRGÅENDE INDSATSOMRÅDER

Bevaringsplanlægning og -metodik

En troværdig digital langtidsbevaring forudsætter grundig bevaringsplanlægning. Ved bevaringsplanlægning forstår vi dels at udarbejde konkrete bevaringsplaner for datastrukturer og filformater, dels mere overordnet at vurdere, hvornår den teknologiske udvikling nødvendiggør migreringer af enten medier eller formater.

Et væsentligt element er vores dialog og samarbejde med dataskabere og øvrige interessenter, vidensarbejde, omverdensanalyser og løbende teknologiovervågning, som dækker aktiviteter ved både indsamling, bevaring og tilgængeliggørelse.

Inddrage brugere i vurderingen af formaters vigtige egenskaber

De formater, som Rigsarkivet udvælger og anviser til digital arkivering, skal kunne repræsentere de vigtige egenskaber i de originale dataformater så godt som muligt. Vi skal også samarbejde med brugere af originale data i vurderingen af, hvilke af formaternes egenskaber der er vigtige at bevare for at sikre, at data kan genanvendes til de ønskede formål.

Rigsarkivet søger løbende at udvide og tilpasse det udvalg af formater, som anvises til langtidsbevaring, til de behov, som den teknologiske udvikling medfører. På den baggrund kan vi - afhængigt af de ressourcer, der er til rådighed - udpege nye formater til langtidsbevaring. Formatvurdering sker så transparent som muligt ud fra vægtede kriterier og baseres på løbende opsamling af viden og erfaringer.

Anvende flere bevaringsniveauer

Der er imidlertid situationer, hvor det ikke er muligt at udpege et egnet format til langtidsbevaring. For at imødegå denne udfordring vil Rigsarkivet arbejde struktureret med forskellige bevaringsniveauer. Et bevaringsniveau kan fx være at bevare data i originalformat, indtil det evt. er muligt at konvertere data til et format, der er velegnet til langtidsbevaring. Dette kan specielt være relevant, hvis en dataskaber skal nedlægges, eller hvis data skal slettes hos dataskaber af hensyn til persondatalovgivningen.

Der kan også være situationer, hvor data ikke hensigtsmæssigt kan flyttes fra dataskaberen. Det kan fx være, hvis der er tale om data, der ikke kan bevares systemafhængigt; hvis det kræver fagspecifik viden at bevare data; eller hvis der er tale om meget store datamængder. I sådanne tilfælde kan data bevares i den kontekst, hvor de er skabt, men være underlagt arkivlovens bestemmelser om bevaring og tilgængelighed.

Skabe grundlag for genanvendelse af originale data

Så vidt det er økonomisk muligt, vil Rigsarkivet arbejde for supplerende at modtage data i de originale formater som et supplement til bevaringsudgaven. Dette kan gøre det lettere at finde konverteringsfejl ved modtagelse og test og samtidig muliggøre, at brugere i en periode kan få adgang til originale data.

Samarbejde og standardisering

Samarbejde om vidensopbygning, metoder og værktøjsudvikling

Arbejdet med at udvikle og vedligeholde løsninger og metoder til digital arkivering er et komplekst område og skal håndteres ved, at bevaringsinstitutioner samarbejder og videndeler.

Anvende fælles europæiske bevaringsformater

Rigsarkivet vil fortsat indgå i og bidrage aktivt til nationalt og internationalt samarbejde, navnlig i regi af E-ARK-projektet, men også på nordisk plan. Vi vil arbejde for at indarbejde de relevante fælleseuropæiske bevaringstandarder i Rigsarkivets regler for digital arkivering. Dette vil forbedre myndighedernes muligheder for at konkurrenceudsætte arkiveringsopgaver og Rigsarkivet vil kunne udvikle metoder og værktøjer i samarbejde med andre europæiske arkiver. Rigsarkivet anvender OAIS-modellen som referenceramme for arkivets interne processer for digital arkivering.

Det er en høj prioritet for Rigsarkivet at samarbejde med vores interessenter i forhold til at videreudvikle og vedligeholde standarder, regler, vejledningsmateriale osv. Dette gælder også i forhold til ikke-statslige og private data. Digital arkivering indgår allerede som en vigtig del af myndighedernes informationshåndtering på samme måde som fx databeskyttelseslovgivningen, offentlighedslov, forvaltningslov og god forvaltningsskik. Rigsarkivet arbejder for at sikre, at der er et så hensigtsmæssigt samspil som muligt mellem både de tilgrænsende lovgivninger og standardiseringsinitiativer og de ansvarlige myndigheder.

Troværdige data

Dokumentere processer udført på data i hele deres livscyklus

Rigsarkivet har som mål at bevare digitale arkivalier på en måde, der sikrer, at deres troværdighed oprettholdes. Troværdighed er som udgangspunkt forankret i dokumentationen af de processer, som data har været igennem.

Det er Rigsarkivets målsætning i samarbejde med myndigheder og leverandører at udbygge sporbarhed og dokumentation ift. de ændringer, som af bevaringsmæssige grunde er sket med data undervejs fra et it-system til arkiv. Denne sikring og dokumentation af datas autenticitet kan fx ske under afleveringen ved at sammenligne de arkiverede data med data i driftssystemet. Desuden er der fokus på at understøtte sporbarheden for data efter modtagelse på arkivet i forbindelse med test, bevaring, genanvendelse og fremtidige migreringer ved hjælp af tjeksummer, dokumenterede procedurer, konverteringsnotater m.v.

Innovation

Analysere mulighed for løbende aflevering

Rigsarkivet vil skabe rum for eksperimenter, der kan fremme hurtig implementering af innovative tiltag i digital bevaring. Det kunne være analyse af løbende aflevering til arkiv, forskning i at anvende eller bevare data eller bidrag til at udvikle anonymiserings- eller testværktøjer, der anvender kunstig intelligens.

Anvende og udvikle innovative teknologier

Det er en målsætning, at Rigsarkivet til stadighed arbejder på at udvikle, implementere og anvende relevant ny teknologi, som kan understøtte aktiviteter i arkiveringsprocessen, fx automatisering, kunstig intelligens og maskinlæring.



RIGSARKIVET

www.sa.dk