

Langsigtede løsninger skal sikre biblioteksmaterialerne

Statsbiblioteket lægger afstand til kortsigtede løsninger og går i stedet efter langtidssikring af biblioteksmaterialerne. Et fælles system skal klare håndtering, tilgængeliggørelse og langtidsbevaring. Projekterne DOMS og Det Nationale Bitmagasin skal være med til at realisere den strategi.

Af Bjarne Andersen bj@statsbiblioteket.dk, Statsbiblioteket

Statsbiblioteket har udviklet og benyttet systemer til onlineadgang til digitale biblioteksmaterialer helt tilbage fra begyndelsen af 1990'erne, hvor ISDN-forbindelser var noget de færreste havde i hjemmet, og hvor f.eks. Real-Audio var et meget udbredt format til online distribution af lyd. Netop formidling var ofte i fokus, når der blev gennemført pilotprojekter. Det har gennem tiden betydet, at der på Statsbiblioteket er udviklet mange forskellige tjenester med onlineadgang til mange forskellige små digitale samlinger.

De kortsigtede løsninger har præget billedet

På grund af internettets udvikling blev alle disse pilotprojekter samlet set ofte præget af hyppige skift af både bagvedliggende og brugervendte teknologier. Fokus på løsningernes holdbarhed i et længere perspektiv var ofte fraværende, fordi man ville tilbyde biblioteksbrugere nem og hurtig adgang til digitale materialer. Langtidsbevaring af de digitale objekter var heller ikke rigtig et fokusområde for 10 år siden.

Det førte i 2007 til, at Statsbiblioteket besluttede at udvikle en ny digital strategi, hvor først et analyseprojekt og derefter flere implementeringsprojekter skulle gennemføres. Projekterne skulle realisere visionen om et fælles system til håndtering, tilgængeliggørelse og langtidsbevaring af alle bibliotekets egne digitale samlinger. To af disse implementeringsprojekter er DOMS (Digital Object Management System) og Det Nationale Bitmagasin (bitrepository).

Håndtering af adgangsstyring

Statsbibliotekets DOMS-projekt implementerer en generisk infrastruktur til håndtering og bevaring af digitale objekter. Kernen af systemet baserer sig på open source arkiv-systemet Fedora Commons (fedora-commons.org), som er et af de mest udbredte arkivsystemer inden for biblioteker og arkiver i det meste af verden. DOMS håndterer input fra mange forskellige kilder af såvel data som metadata.

DOMS tilbyder adgang til data gennem mange forskellige kanaler – både bibliotekets egne digitale services

(f.eks. Mediestream), men også gennem andre relevante kanaler (f.eks. Europeana). Adgangen til data og metadata styres af et modul, der tilbyder adgangsstyring på mange forskellige niveauer via f.eks. IP-nummer og WAYF. På den måde kan vi afspejle de mange forskellige rettighedsmodeller for det tilgængeliggjorte materiale.

DOMS implementerer også den administrative brugergrænseflade, der skal gøre biblioteksmedarbejdere i stand til at administrere og f.eks. metadatere de digitale samlinger og objekter. DOMS blev sat i drift i en tidlig version i 2009 og i dag er det en helt central komponent i Mediestream.

Bitsene skal bevares intakte

En anden central komponent i den næste generation af vores digitale infrastruktur er det såkaldte bit-magasin (bit repository). Bit-magasinet skal grundlæggende sikre bit-integritet over tid – altså at alle 0'ere og 1'ere forbliver 100% intakte. Det løses typisk gennem flere uafhængige kopier af data samt løbende kontrol af kopierne vha. såkaldte checksummer – en slags fingeraftryk af de enkelte digitale filer. Der kræves ret avanceret og særdeles robust software til at styre disse processer, og softwaren skal i sig selv heller ikke styre for meget, idet det er katastrofalt, hvis den ved en fejl sletter alle kopier. Koncepter og den centrale software udvikler vi i tæt samarbejde med Det Kongelige Bibliotek og Statens Arkiver. Projektet Det Nationale Bitmagasin er finansieret af Kulturministeriet. Fremover forventes samarbejdet at fortsætte omkring blandt andet vedligehold af software.

En af de store tekniske udfordringer i fremtiden er de digitale systemers evne til at skalere til meget store datamængder. Allerede nu bevarer Statsbiblioteket over 1 petabyte data (1000 terabytes), og vi forventer at det vil stige til 5 petabytes over de næste 5-7 år. Netop skalering er hovedfokus i et nyt EU-projekt kaldet SCAPE (Scaleable Preservation Environments), som Statsbiblioteket som eneste danske partner deltager i. 