

Vi åbner kulturarven

- Interview med Ditte Laursen

“Det Kgl. Biblioteks digitale samlinger tilhører hele Danmark, og vi har pligt til og ansvar for ikke bare at bevare dem, men også give adgang til dem, så flere får glæde af dem, til gavn for oplevelse og oplysning”

”Vi oplever højere og højere efterspørgsel på adgang til digitale materialer som åbne data, og vi skal servicere forskellige brugertyper på forskellige måder. Der er studerende og forskere, der foretrækker at arbejde med de digitale samlinger i deres egne værktøjer, som de kender og har brugt til andre projekter. Det stiller udfordringer til os som organisation. Det er en læringsproces for os. En ny måde at arbejde på.”

Ditte Laursen er afdelingsleder for Digital Kulturarv, som har ansvaret for de digitale nationale samlinger på Det Kgl. Bibliotek. REVY har inviteret sig selv på besøg på hendes kontor i Den Sorte Diamant i København. Til en snak om data. Åbne data og den vifte af erfaringer, som hun og hendes kolleger, der arbejder med digital kulturarv, har gjort sig. Både omkring nye brugerbehov og det at åbne og klargøre data, og stille dem til rådighed for søgning, fremvisning, download og data mining.

”Den grundlæggende præmis er, at vi gerne vil gøre mere indhold tilgængelig for flere. Vi har et demokratisk ansvar for, at der er adgang til den kulturarv, der tilhører os alle.”

De åbne data

På Det Kgl. Bibliotek findes en lang række digitale samlinger, som er åbne for alle. Det er mest digitaliserede materialer, der har et fysisk forlæg, fx

bøger, tidsskrifter, småtryk, håndskrifter, fotografier og noder.

”Alt som ikke nyder ophavsret kan vi principielt stille til rådighed som åbne data. Det er mest ældre materialer. Håndskrifter, aviser, og breve. Men fra Netarkivet kan vi også stille web-dokumenter af nyere dato til rådighed. Det som er udgivet af det officielle Danmark, online rapporter i pdf fra ministerier eksempelvis, dem kan vi også stille frit til rådighed. Den største samling er aviser, der har vi digitaliseret 35 millioner avissider, og 6.1 million sider kan downloades af alle. Det er en rig og stor mængde data. Vi ville gerne stille endnu mere til rådighed, men copyright og mangel på ressourcer sætter en dæmper på det.”

”Det der er vores største udfordring med de åbne digitale samlinger, er at skabe nogle ordentlige adgange. De ligger lige nu spredt i forskellige platforme, noget ligger i bibliotekssystemerne og noget skal man finde i andre applikationer. Samtidig er der materialer, som kun er i vores interne systemer, fx fordi de mangler metadatering. Man skal vide, at det er der, og hvor det er, for at kunne finde det. Som bruger er det svært at navigere i”.


Efter fusionen mellem det gamle Kongelige Bibliotek og Statsbiblioteket og indlemmelsen af digitale samlinger fra andre institutioner arbejdes der på nye

fælles indgange til de digitale samlinger, sammen med udarbejdelsen af en plan for oprydning og nedlukning af eksisterende grænseflader.

”Det er et indsatsområde, at vi skal forbedre adgangen til de digitale samlinger. Vi arbejder med at konsolidere og sammenlægge platforme. For eksempel laver vi lige nu en tekst-platform, som skal indeholde digitaliserede bøger, håndskrifter og breve. I stedet for som nu at have det på forskellige platforme. Så med en ny fælles løsning lukker vi efterhånden de op mod 50 platforme, som er opstået i årenes løb, fx på baggrund af midlertidige projektmidler.” Ditte Laursen fortsætter, ”En konsolidering betyder også forenkling, ikke blot med henblik på at gøre det lettere at navigere i for brugerne, men også at vi fremadrettet skal vedligeholde færre systemer.”

Ud over arbejdet med nye brugergrænseflader i Digital Kulturarv arbejder Det Kgl. Bibliotek også med et nyt website til kb.dk, der vil indeholde introduktionstekster og materialeguides til samlingerne.

”Der vil vi også gerne vise, at der er forskellige indgange til samlingerne. På nogle platforme kan man se materialet, søge det frem, se en pdf og bladde i den enkeltvis. Og så er der adgange, hvor man kan bruge et API”.



Et API er en softwaregrænseflade til indholdet, henvendt til softwareudviklere eller tekniske forskere. API'et kan modtage og besvare forespørgsler, fx om at levere tekst fra en bestemt tidsperiode eller kategori.

”Via API'et kan brugeren hente sit materiale som et datasæt og arbejde videre med det på egen hånd, med egne værktøjer og for eksempel lave data-analyser, visualiseringer og samkørsler med andre data. Det er hele Digital Humaniora-perspektivet.”

”API'et giver også andre mulighed for at lave deres egne platforme. Hvis et forlag vil lave en læringsressource, baseret på materialer som vi har, så kan de selv lave det, ved at programmere op imod API'et. Når vi opdaterer eller laver ændringer hos os, vil det også slå igennem hos dem. Det er en måde at give data fri på. Det er ikke længere kun os, der bestemmer, hvordan materialet skal bruges og hvilke sammenhænge det skal indgå i”.

”Uden API ville vi principielt skulle tilgængeliggøre alle vores data, på en måde der imødegår alle de potentielle brugsscenarier og formidlingskontekster, som samlingerne lægger op til – og det ville være et umuligt projekt”, tilføjer Ditte Laursen.

Brugere og de frie datasæt

Det Kgl. Bibliotek har erfaret, at der grundlæggende er 4 typer af brugere af de frie datasæt. Forskere, studerende, app-udviklere og hackere. Med hacker menes her en engageret ulønnet programmør. De fire typer har hver især meget forskellige behov.

”Forskeren er interesseret i et bestemt emne og er ude efter unikke data, der kan skabe ny viden. De vil være kritiske overfor, hvordan data er skabt. Kan vi regne med dem? Hvad er baggrunden for data, hvordan er de behandlet... og så videre.”

”De studerende ønsker ikke nødvendigvis bestemte data, men mere data de kan øve sig på. Vi oplevede, at de havde svært ved at bruge en API. De skal have et trænings sæt, kan man sige. Der er de her data, de er skabt på denne måde, og derfra kan de arbejde med dem.”

”App-udviklerne er ude efter unikke data, der er fagspecifikke og har potentiale i forhold til udvikling af nye produkter og tjenester. For eksempel bliver vores historiske luftfoto afvendt af DinGeo, hvor boligejere kan lære mere om egen bolig og lokalområde ud fra en række offentlige tilgængelige datakilder.”

”Hackeren er en entusiastisk programmør, der gerne vil eksperimentere med data og systemer, de er ikke så kildeskritiske, til gengæld skal data tale til dem, og data skal være sjove og udfordrende at arbejde med.”

”De forskellige brugertyper har forskellige behov og teknik-kompetencer, og de skal derfor også vejledes og serviceres på forskellige måder, og det er en problemstilling, som skal løses via de nye vejledninger til samlingerne.”

Afdeling for Digital Kulturarv har også gode erfaringer med at samarbejde med universitetsbiblioteksområder-

nes datalabs, HumLab, Social Science Lab, og på KUB Nord. Her supporteres studerende og forskere, der gerne vil arbejde med digitale data, digitale analyseværktøjer og digitale metoder. Vejledningen bliver så foretaget af informationsspecialister derfra.

”På KUB Nord lab ville de eksempelvis gerne arbejde med HC Ørstedes breve. De efterspurgte nogle data, og vi klargjorde og leverede disse data til dem. Det sætter samlingerne i spil, og det er i vores interesse, ikke bare at samlingerne tilgængeliggøres, men også at de kompetencer, som det kræver at anvende samlingerne, er til stede blandt forskere og studerende. Informationsspecialisterne er gennem deres vejledning med til at kompetenceudvikle forskere og studerende, fx ved at vise de måder, der kan arbejdes med kilddata på i værktøjsprogrammer.”

”Som digital samlingsafdeling bevarer vi og skaber adgang, men vi kan ikke gå ud og snakke med alle forskere og studerende. De forskellige labs har netværket og kanalerne. Og de har for længst opdaget potentialet i Det Kgl. Biblioteks egne samlinger. HumLab har senest kørt vellykkede datasprints i digitaliseret dansk litteratur.”

Et datasprint er en intensiv periode med arbejde i at raffinere, analysere og visualisere data for at skabe indsigt i et emne.

”Vi har lært meget gennem samarbejdet med HumLab omkring at stille data til rådighed som data, i forbindelse med den feedback, som er kommet fra forskerne og studerende og fra labs-med-

arbejderne. Det var fx via Humlabs datasprint at vi blev opmærksomme på nødvendigheden af kuraterede datasæt og filer i højere opløsning. Det samarbejde udvikler også os, og vi kan lave bedre tjenester. I dag har vi et bedre API på grund af det.”

”Samplingsfolkene har også glæde af det. Omkring Dansk Vestindien og hele digitaliseringsprojektet fra samlingerne gik vi til HumLab og sagde, vi har noget fantastisk materiale omkring Dansk Vestindien, skal vi ikke lave noget sammen. Det er win-win. De har kontakten ud i universitetsmiljøet til brugerne, der gerne vil arbejde med emnet, og vores samlingsafdelinger har viden om samlingerne.”

”Samarbejde mellem kuratorer, informationspecialister, og forskere og studerende, det er et godt match til at sikre at de digitale samlinger til staidighed er relevante til brugernes behov”, forklarer Laursen, men understreger også, at der brug for at bibliotekets egne medarbejdere løbende tilegner sig nye kompetencer for at følge med den digitale udvikling.

I den forbindelse driver Det Kgl. Bibliotek med AU Library i spidsen DEFF-projektet Open Science and Digital Literacy. Projektet handler om at understøtte udbredelsen af kompetencer og digitale redskaber til anvendelsen af åbne data på forskningsbiblioteker. Parallelt opbygges et FFU-netværk for digital humaniora.”

Samtidig er der nogle gange brug for en helt individuel og skræddersyet service.

”Forleden havde vi en forsker, der var interesseret i alle danske tidsskrifter frem til 1840. Vi har digitaliseret omkring 100 titler fra den periode. De kan findes som individuelle pdf'er i bibliotekssystemet, men vedkommende ville gerne have alle bind som en

dataklump, for digitalt at kunne undersøge, hvordan religiøse spørgsmål blev behandlet i den periode. Det er en type servicering, som vi også skal håndtere, typisk med hjælp fra vores it-udviklere. Men det er også en meget ressourcerelevende måde at betjene vores brugere på, og derfor planlægger vi at indføre betaling for den slags særlige services.”

Hackaton

Som tidligere nævnt er en af brugergrupperne af de frie datasæt programører, der udfordrer anvendelsen af data og de etablerede systemer og som derfor går under navnet: hackerne. For at varetage den brugergruppes interesser er Det Kgl. Bibliotek medarrangør af et fast årligt hackathon, Hack4DK, en åben workshop for data-nørder. Det er en begivenhed, som har været afviklet siden 2012 i samarbejde med blandt andre DR, Nationalmuseet, Statens Museum for Kunst, og en række andre institutioner. Her har man stillet kulturarvsdata frit til rådighed for fri leg og eksperimenteren henover en weekend.

”De bruger vores samlinger på anderledes og nye måder. De kan komme med nye produkter eller nye koncepter. Det er typisk prototyper, aldrig færdige produkter, men det kan give en masse ideer og feedback til os selv. Det er inspirerende for os.”, siger Laursen. Det understreger også, at perspektiverne i åbne data er mange. De rækker langt videre end nörd-niveauet og universiteternes forskermiljøer. Og åben og fri data, betyder åben og tilgængelig og til fri benyttelse”

”Men åbne data indeholder også en række problemstillinger. Vi har for eksempel billeder af døde børn i vores samlinger, og fanger der sidder i fængsel. Rent teknisk kan materialet være uden for ophavsret, og vi har derfor ret til at stille det til rådighed, men vi er også nødt til at tage højde for etiske

overvejelser”, Ditte Laursen fortsætter, ”Der kan også ske det, når man sætter data fri, at de bliver anvendt i sammenhænge, som vi ikke selv havde forestillet os. Klippet sammen eller kontekstualiseret på en måde, som Det Kgl. Bibliotek ikke har lyst til at blive associeret med. Men det er ikke op til os at bestemme, hvad kulturarven skal bruges til.”

Danmark set fra luften

Der er også mange muligheder i de åbne digitale samlinger for almindelige borgere.

”Mange slægtsforskere bruger de digitaliserede aviser til at finde materiale om deres fortid, og det samme er tilfældet med vores samling af luftfoto. Her kan brugerne se deres fædrenes gård eller den kirke, hvor oldefar var kordegn”, fortæller Laursen. Luftfotosamlingen på Det Kgl. Bibliotek består af flere ældre særsamlinger. De bliver digitaliseret løbende under projektet ”Danmark set fra Luften”, og store dele af Danmark er allerede dækket.

”Google- og cowi-kort køber vi adgang til, men selve applikationen Danmark set fra Luften er udviklet af Det Kgl. Bibliotek. Brugere i hele Danmark har efterfølgende bidraget med præcise geo-data. I 2017 forbedrede vi data-adgangen og nu har en boligportal integreret de historiske fotos i deres visning. API-adgangen bliver også brugt af Geo Fyn, et fælleskommunalt geodata-selskab, som har også bygget en egen applikation med luftfoto-dataene fra samlingen.”

Fremtiden

Europeana, det europæiske samarbejde mellem arkiver, biblioteker og museer, har lavet en analyse, der illustrerer at det kun er toppen af kranssekagen, der er digitaliseret. I Det Kgl. Bibliotek er det cirka 3% af de fysiske samlinger, og Ditte Laursen anslår, at mange andre europæiske lande ligger på samme niveau.

“Det er en stor udfordring i fremtiden. Der er materialer, som det ikke er relevant at digitalisere, men der også indhold, som ikke kan overleve uden digitalisering, det gælder fx samlinger på magnetiske bånd, der kun har få leveår tilbage. Samtidig er det vanskeligere at få adgang til de åbne samlinger, når de kun ligger i fysisk form. Det gælder så meget desto mere for kulturarv, der har international rækkevidde. Brugen er større på online-samlinger end hvis man fysisk skal ind på Læsesal Vest her i Diamanten eller læsesalen i Aarhus.”

Det Kgl. Bibliotek rummer også en række nyere digitalt fødte samlinger, der stadig nyder ophavsretlig beskyttelse, som fx tv-udsendelser, e-bøger, e-aviser og apps. Disse materialer er også med få undtagelser åbne for den enkelte brugers personlige anvendelse, men adgangen er begrænset til fysisk fremmøde på Det Kgl. Bibliotek. Men gennem licensaftaler med folkebiblioteker og forsknings- og uddannelsesinstitutioner arbejder Det Kgl. Bibliotek for at flere brugere får online adgang til mere digitalt indhold.



Blå bog: Ditte Laursen

Afdelingsleder for Digital Kulturarv, Det Kgl. Bibliotek, med ansvar for digital indsamling, bevaring og adgang til digitale samlinger

2012-2016 Seniorforsker på Statsbiblioteket med opgaveansvar for digitale medie- og webarkiver, metadata og digital forskningsinfrastruktur

2009-2012 Post.doc. DREAM, Danish Research Centre for Advanced Media Materials, SDU i samarbejde med SMK, Arken, Mediemuseet, Eksperimentarium og Statsbiblioteket

2007-2008 Forsker på Statsbiblioteket med ansvar for webarkivering, radio/tv og fagreferencen i medier- og kommunikation

1998-2007 Ph.d. i digital interaktion, SDU, og ansættelser ved AAU, AU, SDU og Université Marc Bloch, Strasbourg.

Uddannet cand.mag. fra Aarhus Universitet 1998 i nordisk sprog og litteratur, og musik.

Forfatter til en lang række internationale publikationer om bl.a. social interaktion i og omkring digitale medier, webarkivering, digitalisering, og brugerinteraktion med udstillingsmedier og digitale arkiver.
www.dittelaursen.dk