

**Jan Oldervoll:**

## **Digitalarkivet – eit resultat av lukke eller forstand?**

Sjølv om eg er del av Digitalarkivet må eg vel likevel ha lov til å sei at Digitalarkivet er ein braksuksess. Det er ei teneste som få eller ingen har make til internasjonalt og som har opna arkiva for tusenvis av menneske som elles aldri vil ha kome i kontakt med arkiv eller arkivsaker. Av alle dei som kvar dag sit bøygde over arkivsaker er det berre ein ubetydeleg del om sit på lesesalane i arkiva. Nesten alle sit over skjermene sine. Så kan ein stilla spørsmålet om det er ein meisterplan som ligg bak suksessen eller om det berre vart slik; at det heile kom som ein slags lottogevinst. Eg er kanskje ikkje den best til å sjå på dette, sidan eg har vore med på heile prosessen. Eg er sjølv sagt ikkje ikkje upartisk, men samstundes sit eg vel inne med kunnskapar om prosessen som ikkje mange har. Så eg skal prøva. Men denne vesle artikkelen skal du helst lesa som ei slags minnebok. *Digitalarkivet, according to Jan*, for å vri på John Irving sin boktittel.

For å få til eit slikt prosjekt trengst kapital. Mykje kapital, Hadde denne kapitalen måtta hentast i form av pengar hadde Digitalarkivet ikkje eksistert. Men kapital kan erstattast av kompetanse, arbeidslyst, entusiasme og pågangsmot. Det er her Digitalarkivet har hatt kapitalen sin.

Eit Digitalarkiv treng digitale kjelder, eller eigentleg digitale kopiar eller avskrifter av kjelder. Ein skulle då sjølv sagt byrja med mormonar-kyrkja som på 50-talet mikrofilma kyrkjebøkene som er no lagde ut som digitale bilete i Digitalarkivet, Korleis det gjekk til veit eg ikkje heilt, men eg veit me skal vera takksame både for at det vart gjort og at Riksarkivet insisterte på at originalfilmen skulle bli verande att i Norge.

Eg vil heller børja med 1801-teljinga. Den som i realiteten har ansvaret for at denne teljing finst i ein digital versjon er no avdøde professor Knut Mykland ved dåverande Historisk institutt, Universitetet i Bergen. Om det var på ein kafé han sat, ein gong i siste halvdel av sekstialet, skal eg ikkje ha sagt. Men ha sat saman med direktøren i Statistisk Sentralbyrå, Petter Jakob Bjerve. I alle høve klaga han til direktøren over at det fanst så dårleg med tabellar basert på den første nominelle folketeljinga i Norge, 1801 teljinga. Her burde byrået gjera ein innsats til stor hjelp og nytte for det historisk forskingsmiljøet i Norge og utarbeida eit nytt tabellverk basert på denne teljinga. Direktøren sa straks ja. Men då måtte Mykland syta for at det vart laga eit magnetband dei kunne bruka for å produsera tabellen. Knut Mykland var glad for svaret. Lite visste han at det var SSB som burde vera glad for Mykland sitt svar. Like lite visste han kva han gjekk til. At det gjekk rimeleg bra var nok meir utslag av lukke enn forstand.

Det gjorde heller ikkje eg, då eg ein novemberdag i mitt første semester på hovudfag fekk vitjing av rettleiareren min, Mykland, på lesesalen. Han spurte om eg var interessert i halv jobb for å ta meg av 1801-prosjektet. Kvifor spurte han meg? Han visste her trongst EDB-kompetanse for å kunna gjennomføra prosjektet. Det fanst det ikkje ved instituttet. Heller ikkje eg hadde det, men sidan eg hadde matematikk som fag meinte han eg måtte kunna læra meg det som skulle til. At det var forstanden som fekk Mykland til å gå til meg vil eg nekta for, om det var ei lukke får andre avgjera.

Så sette me i gang. Det første var å skaffa midlar. Ikkje visste me kva som skulle gjerast og ikkje kor mykje pengar som trongst. Men Mykland hadde forskingspolitisk erfaring og meinte at det maksimale me ville kunna få frå forskingsrådet var 10 årsverk. Så søkte me om det, og fekk det. Men fort oppdaga me at me måte ha noko å skriva på. Dette var før PC-en sine dagar. Det meste av EDB-kommunikasjon føregjekk via holkort. Det meinte me var for tregt. Men det fanst noko som heitte holband punchmaskin, av merke Dura. Det var ein slags skrivemaskin som i tillegg til å produsera eit papirark også produserte eit papirband med holkombinasjonar. Men ein slik maskin var dyr, ca 50.000 kr, eller kanskje 500.000 kr i dag. Mykland og eg reiste til Oslo. Direktøren i forskingsrådet sat i budsjettmøte, men kom ut og hørde på oss og gjekk inn att. Ein time eller to seinare kom han ut att og sa at dei hadde fjerna to stipendstillingar og gjeve oss Duraen vår. Om det her var lukka eller forstanden som slo til, skal vera usagt.

Så fekk me etter annonse i avisa tak i ei ung jente som heitte Målfrid Jakobsen. Ho fekk ein bunke kopiar av folketeljinga, Duraen, ein pult og så sette ho i gang. Me synest ho skreiv raskt, men då me etter eit halvår fann ut at me knapt ville bli halvferdig med våre ti årsverk. Noko måtte gjerast. Eg sat nokre dagar og kika på fingrane hennar. Etter det vart me samde om kva som kunne gjerast for å få opp farten. Deretter puncha ho ca 1000 personar om dagen, 200.000 personar om året. Berre Yngve Nedrebø på sitt beste har kunna konkurrera. Alle seinare registreringsprosjekt har lege langt bak Målfrid Jakobsen i effektivitet, langt bak. Etter ti årsverk hadde me registrert, korrekturlese, kontrollert og koda teljinga. Det vart laga ein datafil med talkodar for alle kategoriar som skulle inn i SSB sine tabellar, og tabellheftet vart gjort tilgjengeleg, for Mykland og alle andre historikarar som var interessert i samfunnet på slutten av 1700-talet, byrjinga av 1800-talet. I dette tilfellet var det nok vår lukke og Målfrid Jakobsen sin forstand som berga oss.

Datafilane med folketeljinga var på mange måtar eit avfallprodukt, noko me brukte for å laga tabellfilane. Dei har seinare synt seg å vera dei nyttigast, og var, som mange vil vita, utgangspunktet for Digitalarkivet. Hadde me rundt 1970 vore meir kunnskapsrike enn me var, ville dei ikkje ha eksisterte. På seint sekstital og syttitalet vart det laga digitale utgåver av folketeljingar mange stader, helst for mindre område. Det skjedd ved at historikarar manuelt bestemte ein kode for kvar person for dei variablane dei hadde bruk for. Desse vart så puncha inn i datamaskinen for analyse, på holkort. Me tenkte også litt på å gjera det slik, men rygga attende fordi me var usikre på kva kodar me skulle velja. Me viste for lite om kva som fanst i teljinga. I vår store ukunne bestemte me oss for at me skulle skriva av teljinga ord for ord og i neste omgang bruka datamaskinen til å laga kodar. Her fanst EDB-folk som fortalte oss at det fanst datamaskinprogram som ville kunna gjera dette. Me trudde dei og sette i gang. Men dei eksisterande datamaskinprogram kunne ikkje gjera jobben, men me fann løysingar. Men her var nok lukka mykje større enn forstanden. Men midt på syttitalet sat me altså med datafilar som skulle visa seg å vera svært så nyttige meir enn 20 år seinare.

På Statsarkivet i Bergen sat statsarkivar Egil Øvrebø og var interessert i kva me gjorde. Han fekk lurt oss til eit samarbeid som både dei og me er glad for kom i gang. Dei ville nytta datamaskinen ikkje til å laga tabellar, men til å laga register, for at både dei og lesesalsgjestane raskare kunne finna tak i informasjon og også spara protokollane. Det var særleg emigrantprotokollane dei var interessert i. Då Målfrid Jakobsen var ferdig å registra 1801-teljinga börja ho på emigrantprotokollane for Bergen. Det var starten på ein meir enn 30 år gamal tradisjon med å skriva av kjelder ved Statsarkivet i Bergen. Her trur eg det var meir forstanden enn lukka som rådde.

I 1981 var RHD, *Registreringsentral for historisk data* oppretta i Tromsø, med føremål å registrera historiske kjelder. Dei konsentrerte seg i første omgang om folketeljningane frå 1865 og 1900. Eg trur ikkje det er feil å sei at dei held fram i tradisjonen frå 1801-prosjektet. M.a. brukte dei i starten ein del av programvaren utvikla for 1801-prosjektet. RHD og det arbeidet dei har gjort er nok også meir basert på forstand enn lukke.

Det neste store registreringsprosjektet er vanskeleg å vita kva ein skal sei om; var det lukka eller forstanden som rådde? Televerket hadde tidleg på nittitalet svært mange såkalla overtallige. Telefondamene var blitt utkonkurrerte av automattelefonen. Men samstundes kunne dei etter norsk lov ikkje seiast opp fordi Televerket samstundes hadde bruk for mykje folk med IT-kompetanse. Då var dei pliktige til å omskolera dei overtallige. Som rimeleg kan vera såg Televerket mørkt på det. I staden oppretta dei i 1993 TNM (*Televerkets Nye muligheter*), eit firma som rett og slett prøvde å finna anna sysselsetjing til overtallige. Der kom *Teleslekt* inn i biletet. Televerket oppretta eit firma som skulle registrera folketeljingsmateriale der folk så skulle betala for å søka i, enten ved å ringja inn eller ved søkja på Internett. Prosjektet hadde klart samband med tidlegare registreringsprosjektet ved at folk både frå RHD og 1801-prosjektet vart trekte inn som rådgjevarar. Det vart også brukt programvare frå 1801-prosjektet til registreringa. Her kan ein kanskje sei at forstanden rådde. Men den kommersielle delen svikta; *Teleslekt* vart lagt ned i 1997. Digitalarkivet overtok dataene. Det skal me koma attende til.

Ved ein kombinasjon av lukke og forstand eksisterte det i 1997 store mengder folketeljingsdata i Norge. Om det var lukka eller forstanden som var viktigast kan ein diskutera. Men det var i og for seg ikkje nok til å etablera Digitalarkivet. Ein trong også kompetanse. Den kunne enten kjøpast eller finnast. *Teleslekt* hadde vist at det på 1990-talet knapt fanst ein kommersiell basis for eit Digitalarkiv. Arkivverket rådde ikkje over naudsynte midlar. Ein måtte prøva å finna kompetanse innan miljøet som var interessert i dataene. Då er det på tide å venda attende til 1801-prosjektet.

Då 1801-prosjektet vart starta fanst det i realiteten ikkje noko programvare som kunne brukast. Det fanst ein papirbandlesar på ein datamaskin laga ved marinen i Horten som kunne brukast til å lesa papirbanda (ca 300 km av dei), men programmet for å kunna lagra dei på datamaskinen måtte skrivast. Til og med editorar for å retta i data måtte skrivast. Men det store problemet var programmet for automatisk å koda t.d. yrkestermane i teljinga til dei vel 300 forskjellige yrkeskodane me bestemte oss for å nytta. Her trongst forholdsvis høg EDB-kompetanse. Pengar til å kjøpa denne kompetansen fanst ikkje i 1801-prosjektet. Redninga vart NAVF (Norges Almenvitenskapelige Forskningsråd).

NAVF delte i 1968/69 ut såkalla programmeringsstipend i humanistiske fag. Eg synest å minnast at den noverande riksarkivaren fekk eit slikt stipend. Det same gjorde eg. Stipendperioden vart brukt til å leia 1801-prosjektet og først av alt utvikla programvare til prosjektet. Det høyrerest sjølv sagt ut som sjølvskryt, men ein del av denne programvaren var faktisk avansert. Men så kan ein spørja korleis det kunne gå bra når ein professor gjekk inn på ein lesesal og nærast talte på fingrane for å finna ein person til å gjera jobben. Det var som den gamle metoden for å læra folk å symja; kasta dei til sjøss og berre dra dei opp rett før dei drukna. Slik lærte eg å symja og slikt lærte eg å programmara. Men det krovs nokon til å dra deg på land, når det trongst. I programmeringa var det no avdøde Mike Gillow, som på 60- og 70-talet var førsteamanuensis ved Universitetet i Bergen sin EDB-avdelinga. Han hadde ei eineståande evne til å inspirera, og til ikkje å la seg stogga av det vanskelege. Han nekta konsekvent å løysa dei vanskelege problema, men var til ei kvar tid viljug til å peika i rett

retning. Han drog ein på land når trongst. For så å kasta ein uti med det same. Eg vart kasta uti mange gongar og lærte å symja. Men eg trur nok at lukka også her var betre enn forstanden. Det var ein kombinasjon av mange tilfeldige treff som gjorde at det gjekk bra.

Tabellane var ferdig, men sjølv synest eg det kunne vera morosamt også å kunna analysera data på ein datamaskin. Men problemet var det var mykje data og den einaste datamaskinen me kunne bruka var UiB sin sentrale datamaskin, med stor trafikk og klar nedprioritering av ressurskrevjande oppgåver. Då måtte eg ta i bruk kunnskapar eg hadde tileigna meg gjennom programmering i første fase av 1801-prosjektet. Resultatet vart at me fekk eit program som kjørte analysar av folketeljinga frå 1801 om lag 1000 gongar raskare enn dei eksisterande statistikkprogramma. Men problemet var at svært få historikarar brukte stormaskiner i det heile tatt. Så kom dei personlege datamaskinane. Ved Historisk institutt først i form av Tiki, som var maskiner med svært liten kapasitet. Teknologien brukt på stormaskiner vart omforma til bruk på desse 'leikemaskinene' og etter kvart på PC-ar. Det me no hadde var altså teknologi som var i stand til å analysera folketeljing på ein svært effektiv måte.

Så kom World Wide Web. Til Norge skjedde det eigentleg i 1994, i samband med OL. Resultat vart då lagde ut på webben, som me kalla det i det dagar. Men også andre ting kom på web, mellom anna telefonkatalogen til UiB. Hans Morten Kind ved IT-avdelinga til UiB laga nokre Pearl-script som gjorde at me fekk fram telefonnummeret til kollegaane våre på skjermen. Eg kika på dei og fann ut at eg burde kunna laga noko tilsvarande for å søka i folketeljingsmateriale. Med god hjelp frå Kind var det i 1995 mogeleg å søka i 1801-teljinga på Internett. Men det var ein lite effektiv måte å gjera det på. Så fann eg ut at eg kunne bruka vanlege Pascal-program til å presentera ting på Internett. Ein sein kveld tok eg fram programma eg tidlegare hadde laga for å analysera 1801-teljinga og tilsvarande materiale på PC, fjerna heile brukargrensesnittet og laga HTML-grensesnitt i staden. Det viste seg å gå bra og 1801-tejinga var på Internett same natta, i 1995.

Dermed hadde me data, me hadde teknologi og med hadde kunnskapar. Men framleis eksisterte ikkje Digitalarkivet. Det var det med arbeidslyst, entusiasme og pågangsmot som ar nemnt i innleiinga. Her kjem først og fremst statsarkivar Yngve Nedrebø inn i biletet. Han har elles namnet sitt frå same bøen i Jølster som den førre statsarkivaren. Finst det ein bø også i midten til å ta over når han trekkjer seg attende? Me har svært gode erfaringar med folk frå den bøen. I tillegg til å sitja med solid arkiverfaringar har han i fullt mon også dei nemnde eigenskapane. I tillegg har begge dei to statsarkivarane sett det som ei av sine hovudoppgåver å gje folk god tilgang til det dei har i arkivet sitt. I 1997, om ikkje før, såg han at Historisk institutt sin teknologi kunne brukast til noko meir enn 1801-teljinga. Han og eg börja drøyma, saman. Me såg også for oss at me kunne etablera eit fellesprosjekt, Digitalarkivet.

Så måtte også Teleslekt pakka saman, i 1997. Då såg me for oss at me kunne overta databasane deira og leggja dei ut på nettet med vår teknologi. For å få det til måtte me ha med oss riksarkivaren, John Herstad. Eg trur nok han møtte ein del motstand i eigen institusjon. Mange såg nok ikkje for seg at Arkivverket skulle driva med slike ting, og særleg ikkje at dei skulle gjera det i samarbeid med Historisk institutt, eller for det eg veit, med meg. Men han fekk departementet med på at me skulle overta både data og også ein del av dei tilsette ved Teleslekt. Dei vart til registreringseiningane våre i Stavanger og på Voss, som no nærmar seg ferdige med å registra 1910-teljinga. Dermed vart Digitalarkivet etablert, 23. Januar 1998, i første omgang som eit samarbeidsprosjekt mellom Historisk institutt, Universitetet i Bergen og Arkivverket, ved Statsarkivet i Bergen.

Dei som skal ha æra for at det vart noko Digitalarkivet er først Yngve Nedrebø, som på den eine sida hadde visjon, og på den andre sida hadde evna til å få med seg dei ha trong for at visjonen skule bli røyndom. Den andre på lista er John Herstad som let seg verta oppglødd av Yngve sin visjon og hadde kløkt til å nedkjempa motstand i eigen organisasjon og også oppnå støtte hos dei styrande. Når det galdt å byggja opp data og kompetanse dreidde seg like mykje om lukke som forstand. Men når det galdt å samla data og kompetanse til ein utradisjonell, men velfungerande organisasjon var det forstanden som rådde.

Kva så med utviklinga gjennom dei ti åra Digitalarkivet har eksistert? Mange har gjort ein stor innsats. Av dei tilsette må først og fremst Anette Skogseth Clausen nemnast. Ho har i det daglege styrt Digitalarkivet, men både kløkt og forstand. Ho har særleg vist evna til å kombinera stor entusiasme for prosjektet med ein utprega realisme for kva som går an å få til og kva som er umogeleg innan dei grensene me opererer.

Men det er også ein annan, Brukaren. Kor mange ulike brukarar me har hatt er det ingen som veit. Det er mange, mange tusen. Somme kjenner me ved namn, særleg frå debattforuma våre, dei fleste ikkje. Men me har fått utallege tilbakemeldingar, alle positive, somme kritiske. Men dei har hjulpet oss å utvikla Digitalarkivet til det det er i dag. Og ikkje minst har dei hjulpet oss med data. Digitalpensjonatet må vel seiast å ha vore ein stor suksess. Eit stort antall databasar har me fått frå brukarane våre for å leggja ut i systemet vårt. Slik sett ber Digitalarkivet framleis preg av å vera eit gigantisk dugnadsprosjekt innan ein organisasjon som på mange måtar er svært konservativ, som skal vera det, men framleis har funne rom til ein slik innsats. Bra!

Ti år er sjølvsgat ikkje ein høg alder. Men kva skal no skje. Først og fremst meir av det same. Fleire folketeljingar. Det vil koma ei ny kvart tiande år. Så burde me få med oss dei eldre. Først og fremst 1875, men helst også 1891-teljinga. Manntalet frå 1660-åra burde vore rydda opp i og 1701 gjort ferdig. Å få lagt biletfilane frå kyrkjebøkene ut har vore ei storbragd. For meg personleg har det vore ei ekstra stor glede at dette er arbeid som er gjort frå Riksarkivet, sidan eg ofte har hatt kjensla av at staben der har vore negativ til å presentera kjelder på nettet. Eg har sikkert teke feil. Her vil det i åra framover koma mykje meir materiale.

Men det er også viktig å gjera biletfilane søkbare på individnivå. Her ser eg for meg at brukarane våre vil vera med på dugnaden. Difor har me no tilbydt eit registreringssystem slik at dei kan registrera indeksar til kyrkjeboksmateriale. Me skal så tilby søkesystem til indeksane.

Kva bør så Digitalarkivet gjera i framtida? Det rimelegaste ein kan sei her er ikkje stort, sidan bemanning er så katastrofalt dårleg som ho er. Her burde vore fleire tilsette. Men på den andre sida har det aldri hindra Digitalarkivet i å gjera ting.

Noko som bør koma, og som sikkert kjem med det første er søk på tvers av alle kjelder. Då vil ein kunna søkja etter Nils Abrahamsen fødd ca 1880 og få fram opplysningane om slike personar frå alle databasar i Digitalarkivet. Denne tenesta ser eg for meg kjem før sommaren, avhengig av kor mykje tid som kan avsetjast til dette arbeidet.

Per i dag er lagt inn 1316 bilete av gardar og gardsbruk i 1900-teljinga. Denne tenesta kan me lettvent utvida til også å gjelda personar. Det kunne fort bli ei fabelaktig biletsamling. Dette er ei teneste det knapt kostar noko å få i gang. Eg har vilja venta til eg var trygg på at det fanst ei

interesse for å leggja ut bilete. Den har me sett er der, så no er det kanskje på tide å setja i gang.

Men viktigare vil eit lenkeprosjekt vera. Lenking kan gjerast på to nivå. For det første kan det gjerast på bustadnivå slik at når ein er inne på ein gard/bruk kan få all informasjon om denne garden/bruket frå alle teljingar. Det same gjeld sjølvst og så for hus i ein by. Eg trur kanskje det beste ville vera å lenkja opp mot matrikkelen av 1886. Dette er ikkje noko lite prosjekt, men ved å ta i bruk alle friviljuge hjelparar er det overkommeleg.

Så kunne dette utvidast til også å lenkja personar. Ein kunne t.d. ta utgangspunkt i 1900-teljinga. Dei personane som står der finst også i ei rekkje andre kjelder. Ein kunne då tenkja seg at det var råd å få opp lenkje til dei saman med 1900-teljinga. Det er fleire måtar å gjera dette på. Ein måte vil vera at brukarar legg inn lenkje i 1900-teljinga (eller andre teljingar) når dei finn ein person nokon annan stad. Med den energien brukarane våre legg for dagen ville det ikkje gå lang tid før me hadde ei anseeleg mengd lenkjer. Men ein kunne også tenkja seg at det vart utvikla eit datamaskinprogram som grovlenkja materialet og at brukarane fjerna lenkjer som ikkje kunne stemma. Men uansett vil dette vera eit prosjekt som både krev ein god del tenking og også mykje programmering. Men samstundes ville det vera eit prosjekt med svært stor nytte både for lek og lærd.

Det siste eg kunne tenkja meg var å kartfesta data. Særleg om ein først hadde lenkja på gardsnivå ville dette vera overkommeleg. Ein kunne då ta utgangspunkt i t.d. matrikkelen av 1886 og kartfesta kvar gard. Det kunne truleg gjerast ut frå offentlege kartdatabasar, om ein fekk tilgang til dei. Dermed kunne ein produsera kart av ymse slag. Men ein kunne også laga samanlikningar over tid på ein mykje enklare måte enn no. Ein kunne t.d. sei at ein ville ha alle personane i databasane som budde i same området som ein gjeven kommune dekkja i 1900. Elles ville ein få nye data. Flytting ville m.a. kunna studerast på ein heilt annan måte enn i dag, fordi ein kunne studera reelle flyttingar også innan ein kommune, og få med flytteavstand. Den kunne bereknast ut frå kartreferansane.

Men når det gjeld nye tiltak og nye funksjonar er det først og fremst auka stab som trengst. Ein skulle tru at ei teneste som på den eine sida er til nytte og glede for så mange tusen menneske og på den andre skaffar Arkivverket eit så godt renommé fortente å få ein større stab enn han har. Det vil vera forstandig av dei som bevilgar, og til lukke for alle oss andre.