

Til vurdering hos SKUPs jury for arbeider i 2014



Av

Tarjei Leer-Salvesen

Atle Brandt

Jan Grytting

Fædrelandsvennen

2. Bidragsytere

Tarjei Leer-Salvesen, epost: tarjei.leer-salvesesen@fvn.no

Atle Brandt, epost: atle.brandt@fvn.no

Jan Grytting, epost: jan.grytting@fvn.no

Redaksjon

Fædrelandsvennen

Tlf 03811

Adresse: Svanedamsveien 10, 5621 Kristiansand

3. Publikasjonsliste

«Sørlandsporten»

Arbeidet ble publisert på www.fvn.no den 3. Juni 2014 og har vært oppdatert ukentlig siden da. Det har også vært lagt til nye tjenester, men det er likevel én dato som bør settes for selve publiseringen av verket.

Nøyaktig adresse: www.fvn.no/sorlandsporten

Andre vedlegg:

Vi legger ved lenker til saker der vi omtaler prosjektet i Fædrelandsvennen og forklarer leserne hva prosjektet er. Dette er et rent nettbasert prosjekt, og det har ingen hensikt å legge ved papirkopier.

3.5.2013: Den brysomme åpenheten (kommentar i forkant)

<http://www.fvn.no/mening/kommentar/Den-brysomme-apenheten-2404107.html>

5.5.2013: Skal være best på åpenhet (nyhetsartikkel i forkant)

<http://www.fvn.no/lokalt/--Skal-vare-best-pa-apenhet-2405142.html>

3.6.2014: Derfor lanserer vi Sørlandsporten (kommentar ved lansering)

<http://www.fvn.no/mening/kommentar/Derfor-lanserer-vi-Sorlandsporten-2612669.html>

3.6.2014: Skryter av Sørlandsporten (nyhetsartikkel etter lansering)

<http://www.fvn.no/lokalt/kristiansand/Skryter-av-Sorlandsporten-2616111.html>

11.8.2014: Sørlandsporten blir større (artikkel ved utvidelse av tilbudet):

<http://www.fvn.no/lokalt/Sorlandsporten-blir-storre-2653905.html>

Redegjørelse for arbeidet

Vi må gi litt bakgrunn for prosjektet for å forklare hvorfor vi ber om at det vurderes av SKUPs jury. Dette er ingen klassisk journalistisk avsløring, og således ikke «et skup» slik man vanligvis oppfatter begrepet. Men Skup er en metodepris. Vi har brukt journalistisk metode som offentlighetsloven og nye dataverktøy for å lage et verktøy og i stedet for en klassisk sak. Siden har det blitt flere saker hvor vi har hatt glede av verktøyet. Men vi vurderer verktøyet som et publisert redaksjonelt produkt, og det brukes av både våre lesere, våre konkurrenter og våre egne journalister. I stedet for å skrive mer om problemet, valgte vi denne gangen å tilby samfunnet en løsning.

Portalen, som drives av Mediehuset Fædrelandsvennen, er et verktøy for undersøkende journalistikk og en service til borgere som trenger åpenhet og innsyn der det offentlige har forsømt å levere. Det er et forsøk fra Fædrelandsvennen på å ta pressens samfunnsansvar på alvor i en digital tidsalder, på en ny og innovativ måte.

Grunnideen er å gjøre offentlige dokumenter lettere tilgjengelig. I dag er de fleste statlige instansenes offentlige journaler let tilgjengelige gjennom Offentlig Elektronisk Postjournal. Men ikke alle. Og ser man på kommuner og offentlige foretak, varierer praksis mye.

Starten på arbeidet: Vi har et problem

Kristiansand kommune var en foregangskommune på innsyn og tilgjengeliggjøring av sine dokumenter med søkbar elektronisk postliste med historikk tilbake til 1999. Fædrelandsvennens journalister var flittige brukere av tjenesten. Plutselig en dag, uten forvarsel, stengte tjenesten i januar 2013. Vi skrev om det vi trodde var en feil, men snart ble det klart at stengningen var resultat av en beslutning i administrasjonen, som ønsket mindre mas med innsynsbegjæringer. Avisen protesterte og møtte kommuneadministrasjonen for å få tilbake et helt sentralt arbeidsverktøy.

Som en midlertidig løsning valgte kommunen å skrive ut postlister daglig, og skanne dem inn igjen som ikke-søkbare bilde-PDFer. Det blir fort mye å lese. Etter en stund lå det 30.000 sider her som må gjennomleses manuelt hver gang man vil finne et brev med et spesielt stikkord. Det er for mye i en hektisk hverdag. Siden bestod av 900 lenker til disse bilde-PDFene.

Ordfører, rådmann, økonomisektor og organisasjonssektor

4. - 8. februar	11. - 15. februar	18. - 22. februar	25. - 26. februar	27. februar
28. februar	1. mars	4. mars	5. mars	7. mars
8. mars	11. mars	12. mars	13. mars	14. mars
15. mars	18. mars	19. mars	20. mars	21. mars
22. mars	25. mars	26. - 27. mars	02. april	03. april
04. april	05. april	08. april	09. april	10. april
11. april	12. april	15. april	16. april	17. april
18. april	19. april	22. april	23. april	24. april
25. april	26. april	29. april 30. april	02. mai	03. mai
06. mai	07. mai	08. mai	10. mai	13. mai
14. mai	15. mai	16. mai	21. mai	22. mai
23. mai	24. mai	27. mai	28. mai	29. mai

Kristiansand Kommune 08.04.2013
Offentlig journal Periode: 05-04-2013 - 05-04-2013

200706448-39 I
Dok.dato: 03.04.2013 Jour.dato: 05.04.2013
Arkivdel: PENSJ1
Tilg.kode: UO Par.: off § 13 jf fv § 13, 1. ledd
Avsender: Kommunal landspensjonskasse- regionskontor Bergen
Sak: Uførepensjon - Avskjermet
Dok: Krav om refusjon - Avskjermet
Saksansv: Kristiansand kommunale pensjonskasse / ANHA
Saksbeh: Kristiansand kommunale pensjonskasse / ANHA

200708841-2 I
Dok.dato: 03.04.2013 Jour.dato: 05.04.2013
Arkivdel: PENSJ1
Tilg.kode: UO Par.: off § 13 jf fv § 13, 1. ledd
Avsender: Kommunal landspensjonskasse- regionskontor Bergen
Sak: Uførepensjon - Avskjermet
Dok: Krav om refusjon - Avskjermet
Saksansv: Kristiansand kommunale pensjonskasse / ANHA
Saksbeh: Kristiansand kommunale pensjonskasse / ANHA

Vi skrev om problemene og involverte leserne. Saken kom opp til politisk behandling, og bedring ble lovet. Men så gikk månedene ute at noe skjedde. Det ble høst og det ble vinter, og arbeidet med innsyn i kommunale prosesser i vår kjernekommune var vrient.

Et verktøy tar form

Så kom journalist Tarjei Leer-Salvesen på en idé. Hvis man tar alle journalsidene og kjører dem gjennom moderne tekstgjenkjenningsprogramvare (OCR), kan vi kanskje lage et søkbart system til intern bruk. Han testet med noen sider i en mappe på en intern server, med godt resultat.

Redaktørene likte tanken, og medieutvikler Atle Brandt ble koblet inn. Etterhvert kom også Jan Grytting inn i arbeidet og bisto med design.

Etter tips fra VG registrerte Fædrelandsvennen seg som brukere av DocumentCloud, en tjeneste utviklet av Centre for Public Integrity i Washington, som skal hjelpe journalister å håndtere store mengder dokumenter på en sikker måte. Tjenesten er gratis for journalister. Alle Kristiansands journaldokumenter ble lastet opp her og lest med OCR-programvaren Tesseract.

Vi samlet alle relevante postjournaler i landsdelen, både de regionale og de for hver enkelt kommune. Noen har søkbare postjournal, andre har ikke. Kommunene ble inndelt i fargekoder etter hvor godt de har tilrettelagt sine offentlige dokumenter. Denne siden ble hovedsiden i portalen «Sørlandsporten».

Atle skrev en egen søkemotor for Kristiansands postjournal, som kun søker i journaldokumentene Tarjei har «reparert» med OCR-programvare og lagt i DocumentCloud.

Sidene som viser resultatene og gjør det mulig å søke er skrevet med AngularJS, et nytt bibliotek som er velegnet for raske, dynamiske og mobilvennlige løsninger. I likhet med øvrige små løsninger på fvn.no benytter vi rammeverket Bootstrap til grunnleggende design av sidemalen, knapper og menyer.

En ulempe er at nettleseren Internet Explorer ikke fungerer sammen med serveroppsettet til DocumentCloud. Til tross for ulike forsøk på å jobbe rundt utfordringene får vi ikke løst dette problemet, og anbefaler folk å bruke en hvilken som helst annen nettleser mens de gjør søk i tjenesten vår.

Slik ser portalen ut: Den kan besøkes på www.fvn.no/sorlandsporten

SØRLANDSPORTEN

Den raskeste veien til innsyn i offentlige dokumenter

Alle kan lese postlistene til kommuner og andre offentlige etater. Vi har linker til postlistene for alle Agder-kommunene på denne siden, sammen med viktige statlige institusjoner og kommunale foretak. Enhver kan be om innsyn i offentlige dokumenter, og det er gratis å gjøre det.

Fædrelandsvennen har laget søkbare postjournal for Kristiansand kommune, Sørlandet Sykehus og for Universitetet i Agder. Disse vil bli oppdatert inntil disse publiserer egne løsninger som er søkbare og har lengre historikk.



Status for digital tilgang på postlister hos de øvrige kommunene

Søkbare med arkiv



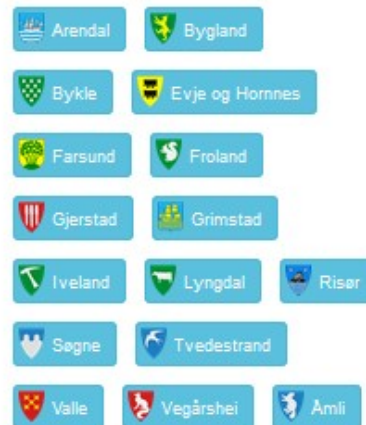
Disse kommunene har digitalt søkbare postlister og digitalt tilgjengelig arkiv og burde være forblide for de øvrige kommunene i Agder.

Uten søkemotor



Disse kommunene har postlister lagt ut som tekst, men mangler søkemotor. Her kan det være lettere å søke via Google.

Søkbare



Disse kommunene har søkbare postlister som fungerer rimelig bra. Knappene fører deg direkte til kommunenes egne sider.

Verktøyet gis bort

Vi hadde fikset problemet vårt, og begynte å lage saker igjen basert på journalsøk slik vi var vant til. Det hadde nå gått et år siden Kristiansand lovet å fikse problemet, uten at det var gjort. I redaksjonen ble vi enige om at problemet er større enn det som angikk oss selv. Kommunen har rett og slett blitt mindre åpen for innsyn for innbyggerne sine. Det er et demokratisk problem.

Da vi besluttet å gi bort verktøyet vi har laget, var det for å bidra til en åpnere landsdel. Vi ønsket å bruke vår innflytelse som mediehus til å presse offentlige etater til å skjerpe seg på åpenhet og innsyn. Gode søkbare postjournal er en offentlig oppgave. Vi tenkte at hvis vi publiserer vår løsning, vil kommunen før eller siden bli nødt til å skaffe seg noe lignende.

3. juni gav vi bort verktøyet til fri benyttelse for lesere og konkurrenter, og fikk god respons på det.

Vi har også delt kildekoden åpent på GitHub, for å spre ideen og legge til rette for at andre gjør det same som oss i sine regioner.

Utvidelse av «Sørlandsporten»

Gjennom sommeren jobbet vi for å utvide verktøyet vi hadde lansert. De to første institusjonene vi utvidet med, var de to andre postjournalene vi savnet mest å ha en god søkefunksjon i, nemlig Sørlandet Sykehus HF og Universitetet i Agder.

Til vår store forundring mente UiA at det fikk holde og publisere tre måneder med offentlig journal og at søkefunksjon ikke kunne være nødvendig. Vi trengte fire årganger for å søke gjennom alt som hadde med den pågående fusjonsprosessen med Høgskolen i Telemark å gjøre.

Det viste seg at alle de åtte universitetene hadde langt mindre åpenhet om sine journaldokumenter enn resten av statlig sektor, som har et utmerket system gjennom Offentlig Elektronisk Postjournal (www.oep.no). Så vi bestilte fire årganger med journaldokumenter fra alle sammen, og laget en tjeneste hvor man kan søke i hvert enkelt av dem, eller i alle universitetene på likt.

Klagesaker etter offentlighetsloven

Mange av universitetene gav oss filer med journaler når vi had om innsyn i dette. Vi sendte ret tog slett innsynsbegjøringer hjemlet i offentlighetsloven og sa at offentlig journal er et document i seg selv. Det var ikke alle som gikk med på det, men hos seks av de åtte universitetene gikk det til slutt greit å få ut det vi trengte.

To universitet har skilt se gut i prosessen: NMBU på Ås og UiT i Tromsø.

NMBU har til en viss grad ikke klart å føre offentlig journal i perioden, og vi måtte klage til Kunnskapsdepartementet og synliggjøre problemene med ressurser til arkivet der. Vi fikk medhold.

UiT var verre. UiT mente først at de slett ikke var nødt til å utlevere offentlig journal ut over de månedene de selv publiserte på sin nettside. To ganger skrev vi klager til Kunnskapsdepartementet om UiT. Begge gangene fikk de fart på seg uten at departementet rakk å realitetsbehandle klagen våre, og omsider fikk vi alt materialet fra dem.

Nå hadde vi mange filer, og en del av PDF-ene våre var mange tusen sider lange lister med dokumentoppføringer. Pro-utgaven av Adobe brukte vi til å stykke opp filene og gi dem navn som lot seg ordne greit kronologisk, slik at databasen sorterer den nyeste filen først.

Et annet problem vi måtte løse, var at bestillingsknappene for dokumentene skal peke til ulike arkivtjenester hos de åtte universitetene og generere en epost for enkel bestilling. Det kunne løses ved å standardisere alle filnavnene, men nå hadde vi lastet ned mange tusen dokumenter med til sammen over en million sider.

Vi fant en løsning, og alle journalfilene er derfor gitt navn med forkortelsene til det aktuelle universitetet.

Filnavnene standardiseres med et lite gratisprogram som heter Bulk Rename Utility.

Fra 30.000 til over en million

Til studiestart utvidet vi portalen. Vi klarte det med enkle tilpasninger av koden Atle Brandt allerede hadde skrevet for vår egenutviklede postjournal for Kristiansand. En litt større jobb var å innhente og bearbeide fire årganger med journaldokumenter fra begge disse institusjonene. Alle filene måtte legges inn kronologisk, og filnavnene måtte fra nå av tilpasses institusjonene for at datasystemene våre skulle holde dem greit fra hverandre.

Det er første gang noe slikt finnes i Norge, en mulighet til å søke i offentlig journal for alle universitetene på likt.

Da vi begynte å laste inn dokumentene fra Kristiansand, ble det 30.000 sider med postlister. Da vi hadde gjort oss ferdige med sykehuset og universitetene, lå det mer enn en million OCR-leste dokumenter på serveren. Nå oppdateres løsningen med mange tusen sider nye dokumenter hver uke.

Vi har også inkludert en søkefunksjon for Sørlandet Sykehus HF som det første helseforetaket i portalen vår. Det er den største arbeidsplassen i vår landsdel som er et offentlig eid foretaket.

Resultater

Det var et uttalt mål med «Sørlandsporten» å ha en oppdragende effekt på offentlig forvaltning på Sørlandet. Vi har ønsket å presse de ulike kommunene og andre instanser til å vise større åpenhet ved å holde høy profil på dette temaet. Prosjektet har skapt debatt og gitt oppmerksomhet slik vi ville.

Kristiansand kommune sluttet umiddelbart i juni å bruke tid på å gjøre PDFene sine til ikke-søkbare bilder da vi viste at vi klarer å reparere dem igjen med OCR-programvare.

I oktober sluttet Kristiansand helt å publisere PDF-lister og lanserte igjen en søkbare postjournal på nett. Denne har bare historikk tilbake til 1. august 2014. For perioden før det må man fortsatt bruke «Sørlandsporten». Vi jobber nå videre med å redde materialet i Kristiansands offentlige journal for perioden 1999-2012 og gjøre det søkbart.

Antallet innsynsbegjæringer steg umiddelbart. De første dagene hadde verktøyet vårt mer enn to tusen brukere. Gjennom høsten lå tallene stabilt på cirka 2000 i uken. Det er uansett langt mer enn bruken internt i Fædrelandsvennen, og vi får tilbakemeldinger om at portalen er i bruk av både konkurrerende redaksjoner, næringsliv og organisasjoner. Det viser at vi har lyktes med å nå ut med prosjektet.

NTNU lanserte søkbar postjournal i november 2014, men bare med tre måneder historikk.

Universitetet i Tromsø hadde søkbar postjournal med tre måneders historikk da vi kontaktet dem i juni. De har nå utvidet denne til ett år, men mener det får holde. Søk med lengre historikk er bare mulig via «Sørlandsporten».

Et annet resultat er at en del arkivmedarbeidere har fått mer å gjøre, og klager til oss over at offentlighetsloven brukes litt vel mye. Vi svarer da at vi håper dyktige arkivarer får flere gode kolleger i fremtiden, for vi håper det skal bli enda mer å gjøre for dem.

Etiske avveininger

Da vi startet med arbeidet ble vi raskt konfrontert med spørsmålet om hvordan vi ivaretar hensynet til personvern. Vi er opptatt av det, og har gjort følgende avgrensinger i hva vi publiserer gjennom Sørlandsporten:

- Offentlige journaler som publiseres her, har vært gjennomgått av etatene for offentliggjøring tidligere. I Kristiansand kommune leser to personer på arkivet gjennom journalene for å fjerne personnavn som ikke skal være der rutinemessig. Det vi gjør, er å gi en lengre historikk og en søkemulighet for journaler etatene selv tidligere har offentliggjort.
- Vår søkbare database er ikke indekserbar mot Google, og derfor vanskeligere å bygge personprofiler med.

Med dette på plass, samt en tydelig melding til brukerne om å si ifra hvis de finner noe feil eller noe som ikke skulle vært publisert, mente vi det var forsvarlig å publisere dette verktøyet åpent.

Kristiansand, 15.1.2015

Tarjei Leer-Salvesen, Atle Brandt og Jan Grytting