

 Oslo 27.06.2017

#  OPPSUMMERING ETTER MØTE ANG EINNSYN

Vi takker for et godt møte hos dere tirsdag 20. juni, der representanter fra Pressens offentlighetsutvalg la fram vår bekymring angående videreutvikling av eInnsyn.

**Bakgrunn:**
Pressens offentlighetsutvalg er representert i referansegruppen som jevnlig tester ny løsning, og juristgruppa som skal diskutere juridiske rammer. Vi har tatt kontakt fordi vi er bekymret for at den nye løsningen for elnnsyn blir så begrenset at det i realiteten ikke blir noen ny løsning, men en videreføring av dagens OEP, med enkelte forbedringer.

Konkret viser vi til at siste versjonutkast legger opp til at man ikke publiserer fulltekstdokumenter direkte i løsningen, men at man i stedet vil ha lenker til etatenes egne nettsider. Dokumentene vil ikke bli indeksert, og det vil derfor ikke være mulig å søke i innholdet i publiserte dokumenter.

Videre vil sperres for direkte nedlasting. Det vil si at systemet vil stoppe deg hvis du får noen tusen treff og forsøker å lese alle sammen ved å klikke deg videre og lese alt. Med andre ord så legges det inn et hinder for at man skal kunne bruke eInnsyn til denne type datastøttet journalistikk.

Etter vårt syn står dette i klar kontrast med hva prosjektet skulle være da det i statsbudsjettet for 2016 ble bevilget 13 millioner kroner til en ny løsning, og kommunal- og moderniseringsminister Jan Tore Sanner uttalte i en pressemelding 7.10.2015 at han vil «styrke demokratiet ved å legge til rette for mer åpenhet og innsyn[[1]](#footnote-1)». Vi viser også til Difis beskrivelse av prosjektet, under tittelen «forbedringer for brukere – innbyggere og presse/media[[2]](#footnote-2):

* *Det kjem****nye og bedre søkefunksjonar****basert på ny søketeknologi og betre datakvalitet*
* *Etter kvart vil det koma****rikare data enn berre datasett****som er definert som offentleg postjournal. Data kan koplast mot andre data om same sak, same sakstype, same geografiske område med meir*

**Begrensninger for søk – publisering av fulltekstdokumenter:**
Da prosjektet startet ble vi forespeilet at eInnsyn skal publisere fulltekst-dokumenter slik mange kommuner allerede gjør. Nå ser det ut til å bli lenker til de enkelte organers mellomlagre i stedet, og dokumentene vil ikke bli indeksert. Dette gjør at det ikke vil bli mulig å søke i teksten i dokumentene. Det innebærer klare begrensninger i muligheten til å søke i offentlige dokumenter. Hvis frykten for søkbare navn er styrende, er vi redde for å miste historikken i OEP når eInnsyn tar over.

Vi spilte inn tidlig i prosessen at søkemulighet i hele dokumenter var en viktig funksjon, noe som er dokumentert i DIfis egen oppsummering i et brukermøte i juni 2016: <https://einnsyn.difi.no/sites/einnsyn/files/160601_presentasjon_referansegurppe_sluttbrukarar.pdf>

Det er begrenset hvor mye informasjon man får inn i en tittel, og vi har mange eksempler på at titler ikke er dekkende for dokumentets innhold. Dette berører også Riksrevisjonen i rapporten Arkivering og åpenhet i statlig sektor, punkt 7.2.1.

**Juridiske begrensninger:**
Utgangspunktet er at man ønsker en løsning som tilrettelegger for mest mulig åpenhet og innsyn fordi det styrker demokratiet. Målet med eInnsyn er å etablere en ny allmenn tjeneste for frivillig fulltekstpublisering, jf. uttalelser fra Jan Tore Sanner i den tidligere nevnte pressemeldingen[[3]](#footnote-3) for Statsbudsjettet for 2016:

«Løsningen vil legge til rette for at dokumenter som etter innsynsbegjæring er vurdert som offentlige, blir publisert i fulltekst. Dette vil bidra til å fjerne tidstyver både for brukerne og ansatte i staten.»

Med et slikt utgangspunkt bør alle begrensinger være godt begrunnet, enten teknisk eller juridisk. Ettersom det teknisk er mulig å lage løsninger med full søkbarhet og nedlastingsmuligheter, må begrensningene ligge i loven.

I dag har vi følgende begrensninger:

a) **Offentlegforskrifta § 6 siste ledd, siste punktum:**«Offentleg elektronisk journal for organ som er omfatta av ordninga etter fyrste ledd, skal innrettast slik at det ikkje skal vere mogleg å få treff på personnamn i innførslar i journalane som er eldre enn eit år.»

**b) Offentlegforskrifta § 7 annet ledd:**«Følgjande skal ikkje gjerast tilgjengeleg på Internett:

a) opplysningar som er underlagde teieplikt i lov eller i medhald av lov

b) opplysningar som det kan gjerast unntak frå innsyn for etter § 9 i forskrifta her

c) opplysningar som nemnde i personopplysningslova § 2 nr. 8

d) fødselsnummer, personnummer og nummer med tilsvarande funksjon

e) opplysningar om lønn og godtgjering til fysiske personar, med unntak for opplysningar om lønn og godtgjering til personar i leiande stillingar i det offentlege og i leiande stillingar eller i styret i sjølvstendige rettssubjekt

f) materiale som ein tredjepart har immaterielle rettar til, med unntak for slikt materiale som er nemnd i § 5 fyrste ledd andre punktum og for materiale der rettshavaren samtykkjer til at materialet blir gjort tilgjengeleg.»

I utkast til forskrift om rett til innsyn i offentleg verksemnd av 04.09.2007 (sak nr. 200705911) presiseres det på side 10 at det ikke skal være opp til hvert enkelt organ å avgjøre hvordan ordningen med elektronisk tilgjengeliggjøring skal utformes, og på side 13 og 14 fremmer Justis- og politidepartementet begrensninger i tilgjengeliggjøringen av dokumenter. Begrensningene har materialisert seg i offentlegforskrifta § 7 annet ledd. Disse setter noen grenser for hva som kan legges ut elektronisk.

**Tekniske løsninger:**

**a)   Teknisk begrensning på å personnavn etter 12 måneder:**

Vi mener det bør være teknisk mulig å legge inn en algoritme som etter 12 måneder (koblet mot dokumentets dato) sperrer mot søk på kombinasjoner av fornavn+etternavn slik de forekommer i Folkeregisteret.

Hvis man kan søke på ordet «nordmann» men systemet sperrer for søk på «Ola Nordmann», så bør man ha ivaretatt de viktigste hensynene når det gjelder forsøk på å automatisere bygging av personprofiler basert på navnesøk. Informasjonen er heller ikke slettet, men det er altså lagt inn en sperre som gjør at man ikke kan søke på kombinasjoner av ord som utgjør kjente, hele personnavn.

b) **Begrensningene i forskriften § 7 (2)**:

Her er det også fullt mulig å sette inn sperringer slik at man overholder kravene i forskriften § 7 (2).

Videre er det klart at når man publiserer fulltekstdokumenter må man sladde slik man gjør ved ordinære innsynskrav, før man legger ut. Man gjør også en forhåndsvurdering av om dokumentet overhodet bør legges ut, jf. «trafikklys-løsningen» som vi snakket om under møtet, og som innebærer at man på forhånd kategoriserer dokumenter med farger som angir om publisering er ok (grønt), trenger bearbeiding (gult) eller er uegnet (rødt).

For forholdet mellom **personopplysningsloven og offentlighetsloven** har Lovavdelingen avgitt en klargjørende uttalelse[[4]](#footnote-4) om dette datert 16.07.2004. Denne gir føringer for hva som bør havne i de ulike kategoriene, men er ikke et argument for en så omfattende begrensning som det her ligger an til.

Videre kan vi ikke se at **kongelig resolusjon av 17.10.2008** om iverksetting av offentleglova og offentlegforskrifta, gir noen begrensninger utover det vi har nevnt over. Tvert imot heter det på s. 35:

«Det vil (…) bli lagt opp til at systemet, så langt det er praktisk mogleg, skal bli innretta slik at det ikkje kan søkjast frå eksterne søkjemotorar i journalane som leggjast ut.»

At det her ikke er tenkt på mellomlagring internt i prosjektet, men eksterne søkermotorer som Google, AltaVista eller Kvasir, bekreftes i Lovavdelings høringsnotat av 2007 nederst på side 12.

Den nye Personvernforordningen artikkel 29 eller for øvrig andre artikler i forordningen, som etter planen skal tre i kraft fra mai neste år, kan etter vår oppfatning heller ikke hjemle de omtalte begrensningene. Vi viser her til forordningens artikkel 85 og 85, der journalistisk virksomhet og offentlighetsprinsippet fremheves. Høringsnotatet fra Justis- og politidepartementet om hvilke virkninger forordningen vil ha på norsk lovgivning, blir etter sigende sendt ut om kort tid. Våre naboland, Sverige og Danmark, har imidlertid allerede slått fast at forordningen ikke krever noen endringer i deres offentlighetslovgivning, jf. SOU 2017:38 «EU:s dataskyddsreform», s. 76 og Justismisteriet «Databeskyttelsesforordningen – og de retlige rammer for dansk lovgivning», Betænkning nr. 1565, s. 959. Vi antar samme konklusjon vil fremkomme i den norske vurderingen. Det er viktig at eInnsyn sees i sammenheng med dette slik at ikke forordningen setter unødvendige begrensninger.

**Faren for personprofiler:**Dersom man vurderer å sette opp begrensninger i søkefunksjonaliteten av hensyn til bygging av personprofiler, etterlyser vi konkrete eksempler på dette.

**Frykten for feil**:
Dersom man innfører begrensninger fordi man frykter feil, mener vi dette er å starte i feil ende. Ønsker man en god løsning må man heller sikre at denne blir så god som mulig ved å legge inn kvalitetskontroll, opplæring mv. i stedet for å begrense løsningen. Feil skjer og vil fortsatt kunne skje, men den erkjennelsen kan ikke danne utgangspunktet for generelle og systematiske begrensninger i en viktig del av den demokratiske infrastrukturen.

**Alternativt - egen løsning for mediene:**
Hvis man velger løsningen som ligger nå, slik at det ikke blir mulig å gjøre egne uttrekk, vil vi på det sterkeste anmode om at denne sperren ikke gjøres gjeldende for mediene. Prinsipielt er vi for at offentlig informasjon gjøres tilgjengelig for alle, men hvis man mener at dette blir for problematisk opp mot personvernhensyn, er en mellomløsning å gi tilgang kun for mediene. Dette vil være i tråd med prinsippet i personopplysningsloven § 7 der flere av bestemmelsene i loven, etter en avveining mellom ytrings- og informasjonsfriheten opp mot personvernet, ikke gjelder for behandling av personopplysninger utelukkende for journalistiske formål. Vi viser særlig til følgende sitater fra forarbeidene til denne bestemmelsen, i Ot. Prp. Nr. 92 (1998-1999, kap. 11.6, som angir hvordan man skal gjøre en slik avveining:

«For den enkelte regel må man vurdere hvor viktig den aktuelle bestemmelsen er for personvernet i forhold til de begrensningene den kan medføre i forhold til utøvelse av ytringsfriheten. I denne sammenheng må man blant annet se hen til om det tas tilstrekkelig hensyn til ytringsfriheten innenfor rammen av den bestemmelsen det er aktuelt å gjøre unntak fra, og om personvernet ivaretas på andre måter enn gjennom lovforslaget (f.eks gjennom andre lovbestemmelser som verner den enkeltes privatliv eller gjennom selvjustisordninger som Pressens Faglige Utvalg og Vær Varsom-plakaten).»

I forarbeidene understrekes også behovet for at pressen får tilgang til sensitive personopplysninger:

«Etter departementets oppfatning kan det ut fra en avveining mellom personvern- og ytringsfrihetshensyn være naturlig også å gjøre unntak fra de generelle behandlingsreglene i § 8, 9 og 11 i lovforslaget, bl.a fordi kravene til saklighet, relevans og formålsangivelse vil måtte strekkes så vidt for den journalistiske virksomhetens vedkommende at det vil være tvilsomt om de vil ha noen selvstendig betydning. Selv med en slik vid utstrekning vil det dessuten kunne reises spørsmål om Datatilsynet bør ha adgang til å avgjøre om f.eks pressens bruk av personopplysninger er relevant og saklig. Sensitive personopplysninger behandles daglig i alle avisredaksjoner og kan ha offentlighetens interesse selv om ikke vilkårene for behandling i § 9 er oppfylt. Som presseorgansisasjonene har pekt på under høringen, har flere viktige saker i norske medier i de senere år omfattet personopplysninger som er sensitive i lovforslagets forstand.»

Og endelig, i likhet med hva som er tilfellet i dag, kan man knytte en plikt til å håndtere opplysningene man samler inn, på en forsvarlig måte i form av sikring og interkontroll:

 «Datatilsynets kontrollvirksomhet og sanksjonsmidler vil være knyttet til de pliktene som loven pålegger de behandlingsansvarlige, og tilsynets tilgang til opplysninger hos de behandlingsansvarlige vil være begrenset på samme måte. Med de begrensningene som foreslås i lovens anvendelse på journalistisk virksomhet, vil det i første rekke være aktuelt for tilsynet å gi pålegg for å sikre at kravene til sikring og internkontroll overholdes.»

Løsningen i eInnsyn må i det minste harmonisere med pressens eksisterende rett til å behandle personopplysninger etter § 7.

Mediene følger Vær Varsom-plakaten og er vant med å ha en særlig tilgang for sensitivt materiale, bla. gjennom domstolenes pressesider der man logger på via BankID. (Mer om kriteriene for tilgang, se her: <http://www.domstol.no/no/Elektronisk-pressemappe/Kriterier-for-godkjenning/>)

Etter det vi forstår er det faren for misbruk og frykt for at det gjøres feil ved utlegg, som er hovedgrunnen til at man vurderer begrensninger. Foruten det vi tidligere har sagt om at dette er å begynne i feil ende, kan det i hvert fall ikke begrunne begrensinger overfor mediene. De presseetiske reglene i Vær Varsom-plakaten stiller strenge krav til publisering av private opplysninger, uavhengig av om man har fått tilgang «rettmessig» eller ved en feil. På de områdene mediene har særlig tilgang i dag, har vi svært få eksempler på misbruk.

**Praktiske konsekvenser – hva risikerer vi å miste?**

Wescraping som journalistisk verktøy er relativt nytt. Samtidig er dette en av de retningene innen datastøttet journalistikk som det nå satses på i utdannelser og kurs og konferanser over hele verden. Slik eInnsyn nå ligger an til å bli, vil det ikke være mulig å laste ned data (webscraping) direkte fra løsningen.

Et norsk eksempel fra 2008, er NRK Brennpunkts kartlegging av branner i verneverdige bygg. Her ble datasett med hvilke adresser det har brent på (med dato), koblet med datasett som forteller hvilke adresser som har fått ulik vernestatus (med dato). Deretter satte redaksjonen opp en webscraper mot Plan- og bygningsetaten i Oslo sin postliste, og hentet ut titler (med gårds- og bruksnummer, samt dato) som inneholdt opplysninger om når gårdeierne hadde søkt om å få rive bygg, og mottatt avslag. Dette gav i sum et kraftig researchverktøy for redaksjonen. Selvsagt ble ikke alle resultatene av spørringene som kunne gjøres her publisert for publikum, det ville være presseetisk uansvarlig. Men for journalistikken var webskraping gull verdt her.

I Norsk Presseforbund fikk vi god nytte av slike verktøy da vi i 2016 hadde behov for å undersøke journalføringspraksis i norske departementer. Et medlem av Presseforbundets Offentlighetsutvalg satte opp en webskraper som hentet inn informasjon om forholdet mellom dokumentdato, journalføringsdato og publiseringsdatoer i OEP. Resultatet var en rapport vi tør håpe dere er kjent med, som avdekket kritikkverdige forhold i en rekke av departementene i forhold til den åpenhetskulturen det er politisk enighet om[[5]](#footnote-5).

Et annet eksempel fra Norge er Innsyn.no, en tjeneste utviklet av redaksjonen Fædrelandsvennen. Her benyttes webskraping for å lage søke postjournaler for offentlige etater som ikke tilbyr dette selv, slik som Forsvaret, politidistriktene og universitetene. Dette er et tilbud som utfyller arbeidet Difi gjør med OEP i dag og planlegger å gjøre med eInnsyn, og som har verdi for oss som journalistisk verktøy.

Videre har vi Kommunal rapports prosjekt om kommunenes leverandører. Journalist Vegard Venli hentet inn store datasett fra landets kommuner og ga informasjonen betydelig merverdi ved å koble den på tvers av kommuner, med data fra Brønnøysundregistrene og skapte en egen tjeneste. Blant sakene som ble til ved hjelp av kopling av persondata, var avsløringer av hvordan flere ordførere fakturerte sin egen kommune fra sine egne privateide firmaer. Data fremskaffet ved hjelp av nettskraping var avgjørende for flere av disse avsløringene. Et eventuelt API er ofte bedre enn både webskraping og manuelle innsynsbegjæringer for denne typen datajournalistikk.

Fra vårt naboland Danmark vil vi trekke spesielt frem firmaet Kaas & Mulvad, som jobber for en rekke danske redaksjoner med opplæring i webskraping og med å skreddersy løsninger for å integrere dette i dansk journalistikk. En av de kjente sakene deres er levekårsanalysene for Fagbladet, hvor de henter statistikk fra en rekke kilder for å oppdatere sine analyseverktøy. Webskraping er metoden som brukes for å holde verktøyene oppdatert.

Et annet eksempel, også fra Danmark, er Berlingske som skrapet data fra auksjonsgiganten Lauritz.com og gjennom dette dokumenterte hvordan bestemte budgivere la inn mange bud, men sjelden kjøpte det de var i budrunde om[[6]](#footnote-6). Det danske Erhvervsstyrelsen tillater API adgang til regnskapsadata og selskapsdata, og det benyttes til masseuttrekk av data.

Vi har også flere eksempler fra andre land som har brukt webskraping til opplysende og viktig journalistikk, se blant annet disse sakene:

*The Guardian:*Gjennom skraping kartla The Guardian hvem som var fakkelbærere for olympiske leker og fant ut at flere samtidig var ansatte av sponsorer for lekene: <https://www.theguardian.com/sport/datablog/2012/jul/11/2012-olympic-torch-relay-torchbearers-sponsors-data>

*BBC and BuzzFeed*: Gjennom skraping av åpne og lukkede odds fra mer enn 26 000 tenniskamper dokumenterte man at mange spillere deltok i spillfiksing: <https://www.buzzfeed.com/heidiblake/the-tennis-racket?utm_term=.qczrKovoa#.esZawOyOL>
*Investico:* Kartla bierverv for nederlandske universitetsprofessorer ved å skrape deres åpne profiler. Fant ut at mye av forskning ved universitetene avhenger av kommersielle interesser:  <http://www.platform-investico.nl/artikel/ondernemende-professoren-nevenfuncties-van-hoogleraren-leiden-tot-belangenverstrengeling/>

Norge har forpliktet seg i Open Government Partnership (OGP) til å lage en løsning som utnytter ny teknologi på en bedre måte både overfor bruker og forvaltning, jf. Punkt 2, B om OEP:

«En ny løsning med bedre kapasitet til å ta imot større mengder data og med ny teknisk løsning som vil gi bedre søkefunksjoner og bedre brukeropplevelser.»

En løsning med søkbare og nedlastbare data, vil gi en langt bedre brukeropplevelse enn det vi nå har fått forespeilet. Vi håper derfor man tar dette på alvor og ser nærmere på hvordan man kan tilrettelegge for at eInnsyn fortsatt blir et solid digitalt verktøy for et velfungerende offentlighetsprinsipp.

Vennlig hilsen

for Pressens Offentlighetsutvalg

   

|  |  |  |
| --- | --- | --- |
| Siri Gedde-Dahl (leder)  | Ina Lindahl Nyrud | Arne Jensen |
| *Journalist i Kapital* | *Advokat Norsk Journalistlag* | *generalsekretær* |

  

|  |  |  |
| --- | --- | --- |
| Øyvind Bye Skille  | Kristine Foss  | Nils E. Øy |
| *Journalist NRK* | *Jurist i Norsk Presseforbund*  | *spesialrådgiver* |

1. <https://www.regjeringen.no/no/aktuelt/ny-oep-gir-enklere-innsyn-mer-apenhet-og-fjerner-tidstyver-i-staten/id2456933/> [↑](#footnote-ref-1)
2. <https://einnsyn.difi.no/om-einnsyn/bakgrunn-einnsyn> [↑](#footnote-ref-2)
3. <https://www.regjeringen.no/no/aktuelt/ny-oep-gir-enklere-innsyn-mer-apenhet-og-fjerner-tidstyver-i-staten/id2456933/> [↑](#footnote-ref-3)
4. <https://www.regjeringen.no/no/dokumenter/-6---avklaring-av-forholdet-mellom-perso/id455092/> [↑](#footnote-ref-4)
5. <http://presse.no/offentlighet/pressens-offentlighetsutvalg/dugnader-og-undersokelser/rapport-departementenes-smks-journalforing/> [↑](#footnote-ref-5)
6. <http://www.business.dk/detailhandel/ti-kunder-staar-bag-millioner-af-bud-paa-lauritz.com> [↑](#footnote-ref-6)