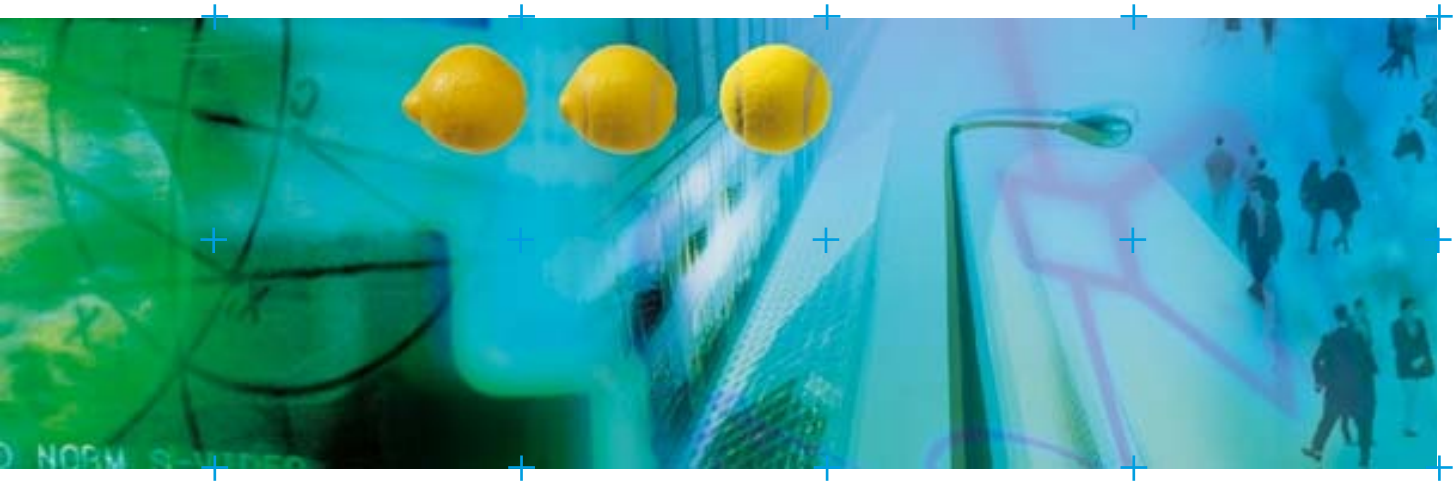
Fra Altinn til alt ut?

Offentlige data for innovasjon og demokrati

Teknologirådet



Rapport 1 2010

Fra Altinn til alt ut?

Offentlige data for innovasjon og demokrati



ISBN 978-82-92447-41-3 (trykt utgave)
ISBN 978-82-92447-42-0 (elektronisk utgave)

Utgitt: Oslo, april 2010
Omslag: Enzo Finger Design AS
Trykk: ILAS GRAFISK
Elektronisk publisert på: www.teknologiradet.no

Innhold

Forord **4**

Sammendrag **5**

Kapittel 1 | Bakgrunn og lovgivning **8**
1.1 Hvorfor er offentlighet viktig? 8
1.2 Teknologiutviklingen 9
1.3 Offentlighetsloven og gjenbruksdirektivet 10
1.3.1 Hvem er omfattet av loven? 10
1.3.2 Opplysninger som er unntatt offentlighet 11
1.3.3 Adgang til å ta betaling for data 11
1.4 Verdien av offentlige data 11
1.4.1 MEPSIR 11
1.4.2 Cambridge rapporten 11
1.5 Ikke-kvantifiserbar verdi 12

Kapittel 2 | Hvilke data? **13**
2.1 Personopplysninger 13
2.1.1 Opplysninger underlagt taushetsplikt 13
2.1.2 Alternativer til å publisere hele datasett 14
2.2 Kvalitet og misbruk 15
2.2.1 Kvalitetsmerking av datasett 15
2.3 Aktuelle datatyper 16
2.3.1 Geodata – kart 16
2.3.2 Statistiske data 17
2.3.3 Forskningsdata 17
2.3.4 Kulturdata 18
2.3.5 Operasjonelle data 18

Kapittel 3 | Forretningsmodeller for offentlige data **19**
3.1 Det offentliges plass i verdikjeden 19
3.2 Pris på data skal være uavhengig av bruk 20
3.3 Databasevern og uforutsigbare betalingsmodeller 21
3.4 Offentlige data bør være gratis 22

Kapittel 4 | En norsk portal for offentlige data **23**
4.1 Hvordan bør portalen se ut? 23
4.2 Ulike formateringsnivåer 23
4.3 Hvem skal forvalte portalen? 24
4.3.1 Finansiering 25
4.4 Hvor skal vi begynne? 25

Litteratur **26**

Forord

Utviklingen innen informasjonsteknologi de siste 10 årene har ført til en eksplosjon i mengden tilgjenglige data – både innelukket i offentlige databaser og i fri flyt på internett. I Norge i dag har innbyggerne i stor grad tatt i bruk PC og internett, og de fleste benytter seg av tjenester som nettbank og selvangivelse på nett. Det har etter hvert vokst fram en forventning hos folk om at man skal kunne finne ut av og ordne det meste på nett. Dette åpner for nye og innovative tjenester basert på offentlige data – men slike tjenester forutsetter at det er mulig å få tilgang til de aktuelle datasettene.

Er det en selvfølge at data som er finansiert av fellesskapet skal være tilgjengelig for alle? Er det greit at noen tjener penger på data de har fått gratis? Hvilken rolle skal det offentlige spille i forhold til kommersielle aktører? Og hvordan kan vi balansere hensynet til åpenhet og hensynet til personvern?

Teknologirådet skal gi uavhengige råd til Stortinget og øvrige myndigheter i viktige teknologispørsmål og dessuten bidra til den offentlige debatten rundt teknologi. Teknologirådet jobber prosjektbasert, og involverer ressurspersoner som har særlig kompetanse innenfor de tema prosjektet omfatter. Bak denne rapporten står en ekspertgruppe med bakgrunn innen ulike fagområder, som jus, offentlig forvaltning, teknologi og IT-ledelse og elektronisk forretningsdrift. Ekspertgruppen har hatt følgende medlemmer:

Espen Andersen – Institutt for strategi og logistikk, Handelshøyskolen BI
Håkon Wium Lie – Opera Software
Jannicke Birkevold – Skatteetaten, Innovasjons- og utviklingsavdelingen
Silvija Seres – Teknologirådet og Microsoft Fast
Tom Slungaard – Norsk Eiendomsinformasjon

Jeg vil rette en stor takk til prosjektleder Christine Hafskjold i Teknologirådet, til ekspertgruppen og til alle andre ressurspersoner vi har fått innspill fra.

Tore Tennøe
Direktør, Teknologirådet

Sammendrag

Fri tilgang til offentlige datasett kan stimulere innovasjon og næringsutvikling og bidra til å styrke demokratiet og kontrollen med forvaltningen. I denne rapporten ser vi på hvilke prinsipper som bør ligge til grunn for tilgjengeliggjøring av offentlige data i Norge. Vi ser blant annet på hvilke data som bør gjøres tilgjengelige for gjenbruk, og hvilke forretningsmodeller som bør legges til grunn for viderebruk av dataene. Vi diskuterer spesielt de personvernutfordringene som er forbundet med offentliggjøring av datasett som inneholder personinformasjon.

Diskusjonen om tilgjengeliggjøring av offentlige data har sitt utspring i EUs gjenbruksdirektiv. Direktivet har som formål å harmonisere hvordan offentlige data gjøres tilgjengelig innen EU og EØS. Tanken er at økt tilgang til slike data vil gi bedre tjenester og økt verdiskapning. Direktivet formulerer krav til hvordan data skal gjøres tilgjengelig, og omtaler også muligheten for å ta betalt for data. I Norge ble gjenbruksdirektivet implementert i Offentlighetsloven 1.1.2009.

Samtidig har konkrete initiativer for gjenbruk av offentlige data i andre land, som USA, Storbritannia og Danmark bidratt med eksempler og erfaringer som kan være nyttige i forbindelse med den praktiske implementeringen av gjenbruksdirektivet i Norge.

Flest mulig datasett bør gjøres tilgjengelig

Som hovedregel bør alle datasett som ikke utgjør en trussel mot rikets sikkerhet eller individers personvern offentliggjøres, ettersom det på forhånd kan være vanskelig å vurdere gjenbruksverdien av et datasett. Data ingen trodde det var interesse for kan få nytt liv i kombinasjon med andre typer data. Gjennom å tilgjengeliggjøre flest mulig datasett kan myndighetene bidra til eksperimentering og innovasjon som kan gi nye og spennende tjenester.

I tillegg til at offentlige data kan komme allmennheten til gode og stimulere innovasjon, kan tilgang til rådata gjøre det enklere å samarbeide og utveksle data på tvers av offentlige etater, og slik bidra til effektiviseringen av offentlig sektor.

Offentlige data bør være gratis

Utviklingen innen lagring og distribusjon av data har ført til at det ikke lenger er nevneverdige kostnader knyttet til disse aktivitetene. Dette medfører at marginalkostnaden (det vil si distribusjonskostnaden) med å dele offentlige data med allmennheten i praksis er null. Hovedkostnaden for offentlige data er knyttet til datainnsamlingen, slik som innrapportering av data fra næringslivet eller innsamling av informasjon om trafikkflyt på hovedveiene. Denne er for de fleste data finansiert av fellesskapet, og tilgangen til dataene bør derfor som utgangspunkt være gratis. Det offentlige tar i dag betalt for enkelte typer data, slik som kart- og eiendomsinformasjon. Gjenbruksdirektivet åpner for å ta betalt, men det er ikke noen *plikt* til å kreve slik betaling.

Det er særlig geodata – og rådata for kart spesielt – som bidrar med direkte inntekter til statskassen. Disse inntektene er imidlertid begrensede i forhold til de offentlige midlene som tilføres Statens kartverk. Lokasjon – hvor noe befinner seg – er en nøkkelreferanse som gjør det enklere å forstå andre data. Kartdata har en spesiell posisjon i forhold til andre typer offentlige datasett, fordi kart kan kombineres med nesten alle andre typer data og bidra til at det blir lettere å forstå dem – det vil for de fleste være mer verd å få byens offentlige toaletter inn på et kart enn å få dem presentert som en liste med adresser.

I dag oppleves prisen på offentlige norske kartdata som en barriere for mindre oppstartsbedrifter. Dette skyldes blant annet at kartdata er priset i forhold til bruk, slik at en videreformidler av kartdata må betale i forhold til antall sluttbrukere.

Generelt er prising etter bruk en modell som gjør det vanskelig for små aktører som arbeider med nettløsninger å forutsi sine kostnader.

Det er i henhold til *databasevernet* anledning til å velge en slik betalingsmodell for informasjon som det er knyttet immaterielle rettigheter til. Det er vår oppfatning at databasevernet kun bør benyttes i tilfeller hvor dataene er tilført et arbeid av betydelig verdi, og hvor vedlikehold og oppdatering utgjør betydelige kostnader for dataeier.

Offentlige data for økt verdiskapning
EU-kommisjonen anslår markedsverdien for offentlige data i Europa til i størrelsesorden 27 milliarder Euro. Selv om tallene i denne undersøkelsen er beheftet med usikkerhet, bekreftes tendensen gjennom rapporter og praksis fra andre land.

Mange offentlige datasett inneholder informasjon som kan bidra til å bedre folks hverdag. Det kan være vanskelig å tallfeste verdien av god tilgang til værvarsel eller togtider, eller av det å kunne se kommunens skolegrenser merket av på et kart. Likevel bør slike ikke-kvantifiserbare gevinster også tas med i verdivurderingen for frie offentlige data.

Det offentliges plass i verdikjeden
Det offentlige bør først og fremst fokusere på å være leverandør av *rådata*. Utvikling av applikasjoner bør i størst mulig grad overlates til markedet, og det offentlige bør unngå å konkurrere med kommersielle aktører. Et viktig unntak er tilgjengeliggjøring av enkeltoppslag på data i de tilfellene hvor publisering av hele datasett utfordrer personvernet, slik som med f.eks. kjøretøyregisteret.

I de tilfellene der det offentlige ser det som riktig og viktig å tilby tjenester rettet mot sluttbruker, bør disse være gratis, og de bør heller ikke eksponere brukeren for reklame. Det er et viktig prinsipp at offentlige tjenester er gratis, og det å bli eksponert for reklame er en form for "betaling" som bør unngås. Det er også viktig at det offentlige ikke blir en konkurrent til private aktører i annonsemarkedet.

Enten dataene er gratis, eller det er fastsatt en pris for dem, bør denne prisen være uavhengig av bruken av dataene. Det offentlige bør heller ikke stille betingelser til bruken eller skille mellom kommersiell og ikke- kommersiell bruk. Dette vil gi en enkel og ryddig forretningsmodell hvor det offentlige ikke bruker ressurser på fakturering og oppfølging av betalingsforpliktelser, men heller kan bruke sine ressurser på kjernevirksomheten sin.

Personvern må settes først
Det er viktig å ta hensyn til personvernet når offentlige data gjøres tilgjengelige. Dette betyr imidlertid ikke at man skal la være å offentliggjøre data: For enkelte datasett kan det være et alternativ å anonymisere dataene. Dette er ikke alltid tilstrekkelig, og det må tas høyde for at personvernet kan utfordres ved at det er vanskelig å opprettholde anonymitet når data utnyttes gjennom nye profilerings- og datautvinningsverktøy. Det er derfor viktig at det gjøres nøye vurderinger før datasett som kan inneholde persondata offentliggjøres, og at Datatilsynet konsulteres i tvilstilfeller.

For datasett som inneholder personlig informasjon, men som ikke er unntatt offentlighet, kan det være aktuelt å gjøre dataene tilgjengelige kun i form av enkeltoppslag. Dette gjør for eksempel Vegvesenet med oppslag i kjøretøyregisteret via en SMS-tjeneste.

Slik vi ser det er ikke hensikten med å tilgjengeliggjøre offentlige data å blottlegge innbyggernes privatliv unødig, men snarere å eksponere det offentlige på godt og vondt – for eksempel i form av data for kvalitet og effektivitet i offentlige tjenester.

Offentliggjøring kan gi bedre kvalitet
Dårlig eller varierende datakvalitet er ingen unnskylding for ikke å publisere datasett. I stedet for at "dårlige" data holdes tilbake, bør de merkes slik at det fremgår hvilken kvalitet de har. Dette gir brukerne mulighet til å ta forholdsregler, slik at data av dårlig kvalitet ikke brukes til for eksempel beslutningsstøtte.

Når datasett blir tilgjengelige, gir det mulighet for tilbakemeldinger fra brukerne dersom de oppdager feil og mangler i datasettene. Slik kan man få en form for digital dugnad for bedre offentlige data. Slike dugnader har tidligere vært brukt med hell, blant annet av yr.no og NRK for å registrere ett koordinatsett pr. postnummer i Norge, slik at det er mulig å få værvarsler knyttet til postnummer.

Data fra offentlig finansiert forskning er også offentlige data

En type data som ikke er omfattet av offentlighetsloven, men som vi likevel vil trekke fram, er forskningsdata. I den grad forskningen i sin helhet er finansiert av det offentlige, bør ikke bare resultatene av forskningen offentliggjøres, men også datagrunnlaget.

En norsk portal for offentlige data

Det bør opprettes en portal som peker til de ulike datasettene hos dataeierne, for eksempel etter mønster fra data.gov i USA. Portalen bør som et minimum offentliggjøre datasett med rådata og programmeringsgrensesnitt (API) for sanntidsdata.

Portalen bør forvaltes av et organ som har ansvar både for kontakt og veiledning av de organene som eier datasettene, og for publikum som ønsker å gjenbruke data. Det bør utarbeides en veileder for dataeiere, og det bør utvikles insentiver for å sikre gjenbruk av flest mulig datasett. Et viktig insentiv kan være at i hvilken grad data gjøres tilgjengelig for gjenbruk blir et kriterium i kvalitetsmålingene til Norge.no.

I utgangspunktet bør flest mulig datasett på plass så raskt som mulig. Enkelte datasett har høyere antatt gjenbruksverdi enn andre, og bør av den grunn prioriteres: Samferdselsdata (informasjon om lokaltransport, trafikk og kø med mer), kartdata og data om kvalitet og effektivitet i offentlige tjenester.

Kapittel 1 | Bakgrunn og lovgivning

Det siste året har det dukket opp portaler for offentlige datasett i flere ulike land: data.gov i USA, digitaliser.dk i Danmark og data.gov.uk i Storbritannia. Portalene er svaret på et økende sug etter offentlige data for ulike formål. Gjennom sin myndighetsutøvelse bruker det offentlige fellesskapets ressurser til å samle inn store mengder data. Teknologiutviklingen og utbredelsen av internett har gjort at stadig flere ser nytteverdien av slike data, og ønsker å få tilgang til dem på en enkel måte, med utgangspunkt i tanken om at offentlige data bør være offentlig tilgjengelige.

Offentlighetsprinsippet forbindes først og fremst med allmennhetens rett til å gjøre seg kjent med opplysninger som kommer frem ved statlige og kommunale myndigheters saksbehandling. Mye av offentlighetsloven og tilhørende forskrifter og veiledere drøfter offentliggjøring av saksdokumenter, hvorvidt dokumenter kan betraktes som *etatsinterne* (for eksempel i en kommune), og hvordan man går fram for å skjerme informasjon som er unntatt fra offentlighet. I dette dokumentet vil vi fokusere på offentlige *data* og *datasett*, og ikke på den typen saksdokumenter som vanligvis forbindes med offentlighetsloven, og som vil tilgjengeliggjøres digitalt i større grad når *Elektronisk postjournal* blir gjort tilgjengelig for allmennheten.

1.1 Hvorfor er offentlighet viktig?

Det er mange grunner til av offentlig informasjon og offentlige data bør være tilgjengelige. De viktigste er hensynet til demokrati og kontroll. I tillegg kan informasjon og data generert i en etat være nyttig for andre, både offentlige og private.

Demokrati og kontroll

Informasjon er en viktig forutsetning for å kunne ta stilling til saker, delta i debatter og påvirke samfunnsutviklingen – viktige prinsipper i et demokrati. I utgangspunktet har tanken vært at innsyn i saker og prosesser gir en mulighet til å kontrollere hvordan byråkrater og politikere forvalter sine posisjoner og samfunnets midler. Dette er noe av bakgrunnen for at pressen i flere år har hatt tilgang til journal- og arkivsystemet til de fleste departementer. Fra 2010 (opprinnelig høsten 2009) skal også allmennheten få tilgang til dette systemet.

Det er imidlertid ikke bare dokumenter som kan bidra til slik kontroll. Bruk av data som en kilde til kritisk journalistikk er på fremmarsj. I Norge har dette vært et av de sterkeste argumentene fra pressen for å opprettholde offentliggjøringen av skattelistene. Gjennom redaksjonell bearbeiding av dataene kan pressen sette søkelys på svakheter i systemet. Man kan også tenke seg muligheten for å sammenholde politikeres programerklæringer og løfter med faktisk stemmegivning i nasjonalforsamling og by- eller kommunestyrer. Et eksempel på en slik tjeneste er www.theyworkforyou.com, hvor man kan få oversikt over medlemmer av parlamentet i Storbritannia, hvordan de har stemt, hvilke uttalelser de har kommet med, hvilke kostnader de har krevd refusjon for, hvilke interesser de har andre steder (annet arbeid, styreverv etc.).

Avisen Guardian har en egen datablogg hvor de gjør tilgjengelig alle datasett de tilrettelegger for eget bruk. Deres erfaring er at mye offentlige data er tilgjengelig på lite fleksible formater, som for eksempel pdf. Gjennom å konvertere dataene til et format der man kan bearbeide dem, har de funnet fram til flere saker som har vakt interesse og debatt.

I Januar 2010 lovet finansminister Sigbjørn Johnsen mer innsyn i departementenes budsjett- og regnskapsinformasjon. Bakgrunnen var kritikk fra Det internasjonale pengefondet (IMF) som pekte på at regnskapstallene er lite tilgjenglige, kommer sent og ikke presenteres på en enhetlig og oversiktlig måte.

Det offentlige som informasjonsbank
I *Offentlighetsmeldingen* fra 97/98 fremheves det at det offentlige er i besittelse av store mengder kvalitetssikret informasjon som kan være en ressurs for private. I NOU om ny offentlighetslov, som kom i 2003, er det mest fokus på den historiske betydningen av dette materialet – at det kan gi et bilde av fortidige prosesser og beslutninger.

I den senere tid har man sett at det offentlige ikke bare besitter store mengder dokumenter, men også datasett, statistisk informasjon, geodata (kart) og lignende som kan være av interesse for private aktører. Både i EUs gjenbruksdirektiv[1] og i Stortingsmeldingen om innovasjon[2] fremheves det at tilgang til denne type data kan ha en positiv effekt på innovasjon, og at det derfor er ønskelig å gjøre tilgangen til slike data enklere.

Likevel kan det være vanskelig å peke på klare bruksområder for alle datasett. I noen tilfeller er bruksverdien åpenbar, slik Meteorologisk institutt har vist gjennom å frigi sine data og utvikle tjenesten *yr.no*. Men selv for data det er vanskelig å kartlegge behovet for på forhånd, kan offentliggjøring ha en verdi. Informasjon er en type gode som man ofte ikke vet at man trenger før man får den. Gjennom å tilgjengeliggjøre flest mulig datasett kan myndighetene bidra til eksperimentering og innovasjon som kan gi nye og spennende tjenester.

For 20 år siden hadde man begrensede mulighet til slik offentliggjøring. Dette kan man fortsatt finne spor etter i dagens offentlighetslov, hvor det blant annet refereres til avskrifter og betaling for kopier. Teknologiutviklingen har imidlertid endret dette bildet dramatisk.

1.2 Teknologiutviklingen

Elektronisk databehandling (EDB) har vært i bruk i norsk offentlig sektor i mange år, med opprettelsen av det sentrale personregister (senere det sentrale folkeregisteret) i 1964 som en viktig milepæl.

Siden den tid er det skjedde mye med måten vi både lagrer og skaffer oss tilgang til data på. På 1980-tallet begynte vi å få databaser, som gjorde det lettere å hente fram igjen dataene, og etter hvert gjorde bruk av nettverk det mulig å oppdatere informasjon i sentrale registre fra kontorer over hele landet.

Fram til 90-tallet var lagringskapasitet mangelvare, og i utviklingen av offentlige – og andre – systemer ble det lagt vekt på å lagre minst mulig data på mest mulig rasjonell måte. Dette ble ikke minst synliggjort ved årtusenskiftet, da en rekke systemer måtte gjennomgås og endres fordi årstall var angitt med kun to sifre for å spare lagringsplass.

De siste 10-15 årene er det skjedd en revolusjon på disse områdene. Lagringsplass er etter hvert blitt så billig at den nesten kan gis bort: En milepæl i så måte kom da Google innførte 1Gb lagring for brukerne av den gratis e-posttjenesten gmail i 2005. Et par år senere – i 2007 – annonserte Yahoo at alle deres e-postkunder ville få *ubegrenset* lagringsplass.

En av grunnene til at det i dag anses som mer lønnsomt å lagre enn å slette, er at man antar at dataene har en verdi. For 15 år siden ville det vært vanskelig å få noe fornuftig ut av så store mengder data, men den radikale utviklingen i søketeknologien har bidratt til at det er mulig å analysere og hente ut informasjon fra nesten alle typer data – ikke bare strukturerte data som er lagret i en database.

En viktig driver – både for innsamling og lagring av informasjon, og utviklingen av søketeknologi – er utbredelsen av internett. I Norge i dag har ca 3 av 4 bredbåndslinje, og så å si hele befolkningen har mulighet til å koble seg til bredbånd der de bor. Etter hvert som næringsaktører har gjort sine tjenester tilgjengelige på internett har det vokst fram en forventning hos folk om at man skal kunne ordne det meste på nett. Dette har igjen ført til

[1] *Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information*
[2] *Et nyskapende og bærekraftig Norge*. St.meld. nr. 7 (2008-2009)

fokus på å gjøre offentlige tjenester tilgjenglige via internett, både i staten – for eksempel gjennom Skatteetatens og Lånekassens tjenester – og i mer varierende grad i kommunene.

Etter hvert som servicenivået på nett blir stadig bedre, oppstår det en forventning om at alle offentlige tjenester og all offentlig informasjon skal være tilgjenglig. Mens nettbrukere tidligere var takknemlige for de tjenestene som var tilgjengelige, er det nå stadig flere som stiller spørsmål ved den informasjonen man *ikke* finner: Hvilken plass har jeg på ventelista for barnehageplass? Hvilken stasjon er 12.40-toget fra Gjøvik kommet til nå? Har restauranten jeg skal til i kveld vært kontrollert av Mattilsynet? Hvor finner jeg nærmeste offentlige toalett? Svaret på disse og lignende spørsmål finnes kanskje i ulike datasett i ulike systemer i ulike etater – både i stat og kommune. Men hvordan skal vi få tak i disse dataene, og hvem har ansvaret for å gjøre dem tilgjengelige?

1.3 Offentlighetsloven og gjenbruksdirektivet

PSI-direktivet (2003/98/EF) tar for seg gjenbruk av informasjon fra offentlig sektor, og omtales derfor gjerne som *Gjenbruksdirektivet* eller *Viderebruks-direktivet*. PSI står for "Public Sector Information", og direktivet er et ledd i arbeidet med å harmonisere EU-landenes praksis for tilgang til offentlige data. Bakgrunnen er en forventning om at bedre tilgang til offentlige data kan føre til innovasjon og nye tjenester. De enkelte landene som er omfattet av direktivet står fritt til å gå lenger enn det direktivet krever, og slik legge om til mer omfattende gjenbruk av offentlige informasjon og offentlige data.

Etter at direktivet ble vedtatt i 2003, har både Norge og de fleste EU-landene jobbet med å utrede hvordan direktivet skal implementeres nasjonalt. Flere av disse utredningene danner underlag for informasjonen i denne rapporten, blant annet:

- *"Norge digitalt" – et felles fundament for verdiskapning*. St.meld. nr. 30 (2002-2003)
- *Fra bruk til gjenbruk*. Tilrådinger fra interdepartmental arbeidsgruppe til Arbeids- og administrasjonsdepartementet, 2004
- *Om lov om rett til innsyn i dokument i offentleg verksemd (offentleglova)*. Ot.prp. nr. 102 (2004-2005)
- *Utkast til forskrift om rett til innsyn i offentleg verksemd (offentlegforskrifta),* Høringsnotat fra Lovavdelingen i Justis- og Politidepartementet, 2007
- *Om gjennomføring av Inspire-direktivet*. Høringsnotat fra Miljøverndepartementet, 2008
- *Et nyskapende og bærekraftig Norge*. St.meld. nr. 7 (2008-2009)
- *Iverksetting av lov 19. mai 2006 nr.16 om rett til innsyn i dokument i offentleg verksemd og forskrift til lova*. Kongeleg resolusjon 17. oktober 2006
- *Rettleiar til offentleglova*. Justis- og politidepartemetnet, 2010 (endeleg versjon)

I Norge er direktivet innarbeidet i Offentlighets-loven som trådte i kraft i sin nåværende form 1.1.2009.

1.3.1 Hvem er omfattet av loven?

I utgangspunktet er alle offentlige virksomheter omfattet av offentlighetsloven, med unntak av Stortinget, Riksrevisjonen og organ tilknyttet Stortinget. De senere årene er stadig flere offentlige virksomheter blitt såkalte selvstendige rettssubjekter, slik som Posten og Mesta. I utgangspunktet gjelder offentlighetsloven for slike selskaper dersom det offentlige direkte eller indirekte kontrollerer mer enn halvpartene av stemmene i det øverste organet i selskapet.

Det er imidlertid visse unntak: Selskaper som driver næringsvirksomhet i direkte konkurranse med og på samme vilkår som private, er unntatt fra loven. Følgende selskaper er nevnt eksplisitt i lovteksten: Eksportutvalget for fisk, SIVA Eiendom Holding, Gassco, NRK Aktivum, Petoro, Posten Norge, Statkraft og Norsk Tipping.

1.3.2 Opplysninger som er unntatt offentlighet

Det finnes en rekke regler som gir adgang til å gjøre unntak fra offentlighet, for eksempel rikets sikkerhet, intern saksbehandling i departementet, eller dersom offentliggjøring kan lette gjennomføringen av straffbare handlinger og personsensitive opplysninger.

1.3.3 Adgang til å ta betaling for data

Offentlighetsloven og forskriften til offentlighetsloven slår fast at det er lov å kreve betaling for innsyn i offentlige dokumenter i visse tilfeller, for eksempel dersom noen ber om å få tilsendt store fysiske dokumenter, eller fysiske dokumenter på kostbare formater.

Selv om offentlige data er samlet inn for fellesskapets midler, er det ikke gitt at dataene er tilrettelagt for viderebruk. I lovgivningen er det åpent for at det kan kreves betaling for informasjon som blir produsert eller bearbeidet kun for å dekke et behov hos eksterne aktører. Denne betalingen kan ikke overstige kostnadene ved innsamling, produksjon, reproduksjon og formidling av informasjonen. Noen aktører har i tillegg lov til å kreve en *rimelig avkastning* på investeringene. Det er dette siste prinsippet som ligger til grunn for den betalingsmodellen vi har for kart- og eiendomsinformasjon i dag. Denne modellen er vanlig i mange land.

1.4 Verdien av offentlige data

I forbindelse med innføringen av gjenbruksdirektivet var det mye spekulasjon om hvorvidt alle offentlige data nå skulle bli gratis. Som beskrevet i avsnittet over er det fortsatt anledning til å ta betalt for data. Dette er imidlertid en praksis som har vært kritisert av mange – både i Norge og i andre europeiske land. Det har vært hevdet av gratis data vil medføre økt innovasjon, ikke minst fordi inngangsprisen for små, nye aktører blir mye lavere.

1.4.1 MEPSIR

Ettersom gratis data vil medføre reduserte inntekter, har myndighetene i EU ønsket å kartlegge inntektspotensialet for offentlige data: Hvor mye er offentlige data verd på det åpne markedet? Den kanskje mest refererte rapporten i denne sammenhengen er MEPSIR-rapporten (Measuring European Public Sector Information Resources) som kom i 2006 på bestilling fra EU-kommisjonen. Rapporten skulle fungere som en målestokk (benchmark) i forhold til senere evalueringer av implementeringen av gjenbruksdirektivet.

Rapporten konkluderer med at markedet for gjenbruk av offentlige data i Europa er på minst €27 mrd. Tallet er basert på spørreundersøkelser både hos dataeiere (offentlige myndigheter) og virksomheter som gjenbruker offentlige data. Det er relativt få aktører som har svart på undersøkelsen, og det er stort sprik i de tallene de ulike aktørene rapporterer. Estimatet er derfor beheftet med stor usikkerhet.

For Norge har dataeierne (15 stk som har svart) estimert verdien av sine data til 93 millioner kr, mens gjenbrukerne (1 stk) har estimert verdien til 280 millioner kr.

Situasjonen i USA

I USA er det ingen åndsverksbeskyttelse på offentlig informasjon, og det er tillatt med ubegrenset gjenbruk uten å angi kilde. Den valgte kostnadsmodellen er marginalprising (eks at man må betale for fysiske kopier), og digital informasjon er gratis. Offentlige myndigheter oppfordres til å ikke konkurrere med privat sektor gjennom å selv utnytte informasjonen kommersielt.

Kilde: MEPSIR

1.4.2 Cambridge rapporten

I en studie gjennomført ved Cambridge University for blant annet det britiske finansdepartementet framgår det at å tilby data gratis eller til marginalkost maksimerer gjenbruk, og at de sosiale og økonomiske fordelene mer enn veier opp for inntektstapet. I studien har forskerne sammenlignet ulike modeller for prissetting av offentlige data, først og fremst *uformaterte* data (rådata), hos aktører som har anledning til å kreve betaling i dag og som er tenkt å være selvfinansierende i størst

mulig grad (såkalte *trading funds*). De ulike modellene er:

- Profittmaksimering
- Gjennomsnittskost (inndekning av utgifter, inkludert utgifter til dataproduksjon)
- Marginalkost (kostnaden ved distribusjon av dataene)
- Nullkost

For elektronisk distribusjon vil marginalkost og nullkost i praksis være det samme.

Studien har sammenlignet eksisterende regimer som i hovedsak praktiserer gjennomsnittskost med en tenkt situasjon hvor samme virksomhet praktiserer marginalkost/nullkost.

Konklusjonen er at for de fleste rådata, inkludert for eksempel rådata for topografiske kart, vil en marginalkost-modell være mer gunstig for samfunnet totalt.

> Kartverket i Østerrike (BEV) har redusert sine priser med 97% for enkelte datasett. Som et resultat av dette, har etterspørselen eksplodert, i noen tilfeller med 7000%. Det har kommet til nye brukere fra små og mellomstore virksomheter, og det er kommet til brukere fra nye sektorer. Prisreduksjonen er blitt dekket inn gjennom den enorme etterspørselsøkningen, og inntjeningen har derfor vært stabil.
>
> Kilde: EU-kommisjonen[3]

1.5 Ikke-kvantifiserbar verdi

Det er ikke gjort noen omfattende undersøkelser av innovasjonspotensialet i Norge. I Danmark har IT- og telestyrelsen bestilt en rapport om *Innovativ udnyttelse av offentlige data*. Heller ikke her er det gjort noe arbeid for å kvantifisere potensialet, men basert på tall fra Storbritannia anslår man at det er snakk om en verdi på flere hundre millioner danske kroner. For øvrig vises det bare til internasjonale undersøkelser som tyder på "en vesentlig samfunnsmessig verdi."

Det er grunn til å tro at en del av den samfunnsmessige verdien som ligger i offentlige data er vanskelig å kvantifisere. Noen av de datasettene som er offentlig tilgjengelige i Norge i dag, tilbys for eksempel gjennom gratis applikasjoner og nettsteder, slik som *yr.no*, og *Trafikanten*. De genererer dermed ikke noen inntekt for staten, men sannsynligvis har de likevel en stor samfunnsmessig verdi:

- At båtfolk har nøyaktig værinformasjon å forholde seg til medfører sannsynligvis færre lystbåtulykker langs kysten.
- At reisende kan velge optimalt transportmiddel basert på sanntidsinformasjon medfører mer smidig trafikkavvikling, mindre jobbfravær og mindre stress.

Slike effekter er imidlertid nesten umulige å kvantifisere.

I tillegg til at offentlige data kan komme allmennheten til gode og stimulere innovasjon, kan enklere tilang til data fra ulike etater gjøre det enklere å samarbeide og utveksle data på tvers av offentlige etater, og slik bidra til effektiviseringen av offentlig sektor.

[3] *Re-use of Public Sector Information – Review of Directive 2003/98/EC*

Kapittel 2 | Hvilke data?

I utgangspunktet bør alle offentlige data gjøres tilgjengelige så lenge de ikke truer personvernet eller rikets sikkerhet. Ettersom det ikke alltid er åpenbart når personvernet er truet, vil vi i dette kapittelet se nærmere på dette. Vi vil også gå inn på enkelte typer datasett det er knyttet spesielle interesser til.

2.1 Personopplysninger

I henhold til offentlighetsloven skal alle offentlige dokumenter og data gjøres tilgjengelige for allmennheten. Det er imidlertid et viktig unntak – sensitive personopplysninger. Slike opplysninger er underlagt taushetsplikt. I henhold til Personopplysningsloven gjelder dette:

- Rasemessig eller etnisk bakgrunn, eller politisk, filosofisk eller religiøs oppfatning
- At en person har vært mistenkt, siktet, tiltalt eller dømt for en straffbar handling
- Helseforhold
- Seksuelle forhold
- Medlemskap i fagforeninger

2.1.1 Opplysninger underlagt taushetsplikt

Likevel kan det av og til være vanskelig å avgjøre om informasjon eller data skal være underlagt taushetsplikt. Det er nemlig ikke bare personopplysningsloven som avgjør dette, også offentlighetsloven og forvaltningsloven tar opp taushetsplikt og skjerming av opplysninger. For eksempel påbyr offentlighetsloven skjerming av følgende opplysninger, utover det som er nevnt over:

- Lønn og godtgjørelse, med unntak for personer i ledende stillinger i det offentlige.
- Der en tredjepart har immaterielle rettigheter og ikke har samtykket til at materialet gjøres tilgjengelig.

I Fornyings- og administrasjonsdepartementets veileder *For innholdsleverandørene til OEP (Offentlig elektronisk postjournal)* er alle disse lovparagrafene tatt i betraktning, og det gjøres en grundig gjennomgang av reglene for taushetsplikt. Veilederen er først og fremst utarbeidet i forhold til å skjerme (i praksis sladde) informasjon i journaler og arkiverte dokumenter, men kan også fungere som en god veileder i forhold til hvilke data som enten ikke kan tilgjengeliggjøres, eller som må anonymiseres før tilgjengeliggjøring.

Men selv om man tar hensyn til retningslinjene knyttet til taushetsplikt, så oppstår det noe nytt når data publiseres på internett, i forhold til når data finnes i et dokument man kan få tilgang til på et offentlig kontor. Søkbare data, og data på formater som kan manipuleres, åpner for muligheter til aggregering og sammenstilling, noe som igjen kan gi et mer omfattende bilde av enkeltpersoner enn det man har forutsatt da datakildene ble vurdert hver for seg.

Det må tas høyde for at personvernet kan utfordres gjennom at det er vanskelig å opprettholde anonymitet når data utnyttes gjennom nye profilerings- og datautvinningsverktøy. I Storbritannia er det tatt til orde for at dette må adresseres dersom man ikke ønsker å få et tilbakeslag for frigivelse av datasett og informasjon.

Skattelistene er et godt eksempel på hvordan informasjonen har endret karakter i takt med at mulighetene på internett har utviklet seg.

Skattelistene

Mange oppfatter publisering av inntekt, formue og skatt som inn inngrep i sitt personvern. På grunn av dette blir skattelistene gjerne målestokken som andre datasett med personinformasjon settes opp mot: Er for eksempel hvilke karakterer man fikk på videregående skole mer eller mindre sensitive enn skattedataene? Fordi skattelistene er i en slik særstilling, har vi valgt å drøfte dette datasettet spesielt.

Norge har en lang tradisjon for å offentliggjøre innbyggernes skatte-, inntekts- og formuesforhold. Offentliggjøringen er hjemlet i ligningsloven og begrunnet i et ønske om å gi innbyggerne mulighet til å kontrollere at skattesystemet fungerer etter hensikten, for eksempel gjennom å sjekke i hvilken grad andre skatteytere bidrar til fellesskapet. Det som offentliggjøres er: Navn, fødselsår, poststed og postnummer, skattekommune, nettoinntekt, nettoformue og utliknet skatt.

Tradisjonelt har skattelistene vært lagt ut til gjennomsyn på lokale ligningskontor. Pressen har imidlertid hatt en særstilling, og har gjennom flere år mottatt skattelistene i elektronisk form for å kunne bearbeide listene redaksjonelt på en hensiktsmessig måte. Enkelte lokalaviser har også trykt skattelistene i sin helhet.

Fra ca 2000 begynte enkelte nettaviser å legge ut skattelistene i søkbar form. Dette har utviklet seg gjennom årene, slik at søkemulighetene og mulighetene for å gå videre og bearbeide trefflistene stadig er blitt mer avanserte. Det er også en økende trend å koble data fra skattelistene til sosiale nettsteder. Allerede fra 2007 fantes det applikasjoner som viste oversikt over skattedata for alle vennene dine på Facebook. Disse applikasjonene er blitt mer utbredt for hvert år.

I tillegg har skattesøk inntatt mobiltelefonene – både i form av enkle SMS-tjenester (betalt pr oppslag) og mer avanserte applikasjoner for smarttelefoner. VGs skattesøk for iPhone (kr 17) var lenge en av de mest populære applikasjonene i Norge. Denne applikasjonen gir mulighet for å søke, finne ulike topplister, koble mot Facebook, sortere, samt plotte resultatene på kart. I så måte er skattedata et godt eksempel på hvordan det dukker opp nye, kreative tjenester og forretningsmodeller i tilknytning til tilgjengelige offentlige datasett. I dag har imidlertid pressen enerett på denne typen tjenester.

Det har vært og er mye diskusjon om hvorvidt offentliggjøringen av skattedata krenker personvernet. Personvern handler om å beskytte folk fra innsyn i den private sfæren. Fordi personinformasjon er nyttig i så mange sammenhenger, gjøres det som regel en avveining mellom hensynet til personvern og det formålet man ønsker å bruke informasjonen til – som for eksempel samfunnssikkerhet eller effektivisering av forvaltningen. Slike avveininger står sentralt i den europeiske menneskerettighetskonvensjonen (EMK) og omtales gjerne som *proporsjonalitetsprinsippet*. I tilfellet med skattelistene vises det gjerne til ytringsfriheten, og til viktigheten av åpenhet rundt hvordan skattepengene samles inn og hvordan byrdene i samfunnet fordeles.

Teknologirådet har tidligere påpekt at de nye måtene skattedataene blir brukt på medfører en endring i proporsjonaliteten, hvor personvernet er svekket.

Det finnes også andre datasett som inneholder personinformasjon og som ikke er underlagt taushetsplikt i dag. Dette gjelder for eksempel kjøretøyinformasjon og vitnemål fra universiteter og høyskoler. Dette er data hvor enkeltoppslag stort sett er uproblematiske, men hvor man kan tenke seg at hele, søkbare datasett kan utfordre personvernet: Tenk deg en Facebook-applikasjon som henter opp informasjon om inntekt, formue, hvilken bil du har og hva den kostet da den var ny, hvilke karakterer du har fra høyere utdanning, hvilke eiendommer du eier, hva de ble kjøpt for og antatt nåverdi etc. for alle i vennelisten.

Slik vi ser det er ikke hensikten med å offentliggjøre offentlige data å blottlegge innbyggernes privatliv unødig, men snarere å eksponere det offentlige på godt og vondt – for eksempel i form av data for kvalitet og effektivitet i offentlige tjenester.

2.1.2 Alternativer til å publisere hele datasett

Til nå har vi stort sett diskutert offentliggjøringen av hele datasett. I forhold til den type data vi har beskrevet over kan det være hensiktsmessig å gjøre unntak fra dette.

Kjøretøyregisteret er i dag åpent for allmennheten via en SMS-tjeneste. Ved å sende REGNR NNxxx til 2282 vil man motta informasjon om bilmerke, fullt navn på eier, informasjon om forrige og neste EU-godkjenning, informasjon om årsavgift eventuelt er skyldig, samt informasjon om omregistreringsavgift for bilen. Prisen er kr. 3,- pr. melding.

En annen, og lignende, informasjonstype er eiendomsinformasjon fra Grunnboken. Denne informasjonen er åpen for alle, men den ligger ikke fritt tilgjengelig i søkbar form på Internett. Enkeltforespørsler kan rettes pr. telefon til Statens kartverk, mens profesjonelle aktører henvises til Norsk Eiendomsinformasjon as. Det er flere hensyn som har begrunnet at Grunnboken ikke er fritt tilgjengelig i søkbar form på Internett, blant annet sikkerhet og personvern.

Avisen Asker og Bærums budstikke har utviklet en søkbar database over eiendomstransaksjoner. Personvernnemnda har slått fast at denne tjenesten ikke er i strid med personopplysningsloven. Dette skyldes sannsynligvis at personopplysningsloven gjør unntak for *journalistisk virksomhet*. Dette er den samme begrunnelsen som brukes av de mediene som gjør skattelistene søkbare på nett. Stortinget har nå bedt Finansdepartementet vurdere måten skattedata offentliggjøres på. Dette vil kunne få konsekvenser også for andre, sammenlignbare datatyper.

Vegvesenet valgte en løsning med SMS-grensenitt og betaling for oppslag i Kjøretøyregisteret etter anbefaling fra Datatilsynet. Bruken av SMS i stedet for en nettjeneste gjør det mye vanskeligere å drive datafangst på denne typen data. Betalingen bidrar også til å begrense oppslagene, samtidig som prisen er satt så lav at den ikke skal være en barriere for dem med reelle behov. Det kunne også være mulig å tenke seg en nettjeneste, men denne ville da måtte inneholde mekanismer for å begrense datafangst.

Denne type metoder bør vurderes også for andre, lignende datatyper. Slik tilgang begrenser muligheten til å utvikle tjenester basert på dataene, men vil også i langt mindre grad føre til misbruk enn om hele registeret lå fritt tilgjengelig i søkbar form. Dersom man er i tvil om det er problematisk å publisere et datasett ut fra personvernhensyn, bør man konsultere Datatilsynet.

2.2 Kvalitet og misbruk

Mange vil mene at det er rimelig at data som er samlet inn med fellesskapets midler også eies av – og deles med – fellesskapet. Mens det tidligere først og fremst var teknologiske barrierer som vanskeliggjorde tilgangen til offentlige data, er det i dag først og fremst utfordringer knyttet til økonomi og organisasjonskultur.

I forbindelse med prosjektet *Fakta først* gjennomførte Institutt for informasjons- og medievitenskap ved UiB en undersøkelse blant 67 statlige virksomheter knyttet til deres planer for frigivelse av datakilder. En stor andel av virksomhetene (63%) oppga at de planlegger å gjøre flere data tilgjengelige for viderebruk i 2010. Undersøkelsen fikk imidlertid også fram en del barrierer knyttet til det å gjøre data tilgjengelig. De viktigste var:

- Det koster (ressurser) å gjøre data tilgjengelig
- Dataene kan bli misforstått
- Dataene kan bli misbrukt til å spre sensitive personopplysninger

Danmark, som i 2009 etablerte nettstedet digitaliser.dk for å få fortgang i tilgjengeliggjøring av offentlige data har gjort lignende erfaringer.

2.2.1 Kvalitetsmerking av datasett

Den motstanden, eller engstelsen, som vi finner både i danske og norske statlige virksomheter er ikke unaturlig, virksomhetenes natur tatt i betraktning. Offentlig informasjon oppfattes i utgangspunktet som ”autorisert”, og i så måte kan det oppleves som problematisk for en offentlig etat å legge ut datasett man vet inneholder feil og mangler.

Når offentlig ansatte er redde for at slike data kan bli misforstått eller misbrukt, er ikke dette nødvendigvis ubegrunnet. Men i stedet for å holde slike data tilbake, bør de heller gjøres tilgjengelige med en klassifisering som angir kvaliteten.

Vi foreslår at et etableres et klassifiseringssystem for kvaliteten på offentlige data. Systemet kan for eksempel se slik ut:

Eksempel på kvalitetsmerking av datasett

A Kvalitetssikret og oppdatert. Kan brukes for eksempel i automatiserte beslutningsprosesser

B Kvalitetssikret og oppdatert. Kan brukes for automatiserte beslutningsprosesser så lenge disse ikke berører liv og helse

C Kvalitetssikret og oppdatert på statistisk nivå, men med mulighet for enkeltfeil

D I hovedsak korrekt, men av mer historisk interesse

E Kan brukes på aggregert nivå som indikasjon på trender og retninger, men må ikke brukes som beslutningsunderlag

F Dataene er ikke oppdatert og kan være beheftet med betydelige feil

En økt åpenhet om ulike typer data kan også bidra til å heve kvaliteten på dataene. Gjennom at det gis innsyn, kan både presse og folk flest påpeke eventuelle feil i datagrunnlaget, som så kan rettes opp hos dataeieren.

Det offentlige kan også legge ut datasett der de *vet* at det er dårlig datakvalitet og invitere til en *digital dugnad* (crowdsourcing) for å rette opp dataene. Slike digitale dugnader har tidligere vært brukt med hell, blant annet av yr.no og NRK for å registrere *ett* koordinatsett for hvert postnummer i Norge og av VG Nett for å oppdatere navnene på alle kommunestyrerepresentanter i landet.

2.3 Aktuelle datatyper

I utgangspunktet bør flest mulig offentlige data gjøres tilgjengelige. Det er imidlertid noen typer data som åpenbart har større allmenn interesse enn andre, og vi vil her se nærmere på noen av disse.

2.3.1 Geodata – kart

Lokasjon er en nøkkelreferanse som gjør det enklere å forstå andre data – både offentlige og kommersielle. Den raske teknologiutviklingen, særlig innen internettapplikasjoner og mobile applikasjoner har drevet fram en sterk vekst i etterspørselen etter geografisk informasjon i konsumentmarkedet. I Storbritannia utgjør selve dataene bare 20% av sluttbrukermarkedet for geografisk informasjon. Markedet totalt er på £900 millioner[4].

Etterspørselen etter navigasjonsutstyr, som GPS-enheter, er så stor at det årlige salget tredoblet seg fra 2006 til 2007, og dette er igjen antatt å mer enn dobles innen 2012. Markedet for navigasjonsenheter er blant de raskest voksende innen forbrukerelektronikk.[5] Markedet er fortsatt under utvikling, og vil sannsynligvis utvikle seg videre gjennom at man får flere lag av informasjon, slik som trafikkinformasjon, værmeldinger, informasjon om destinasjonen med mer som legges "oppå" kartene.

Geodata er kanskje den datatypen som har fått størst oppmerksomhet i forbindelse med gjenbruksdirektivet, fordi kart er den typen offentlig informasjon det er lettest å se gode, kommersielt lønnsomme anvendelsesområder for.

Geodata forvaltes i dag av Norge digitalt, som består av blant annet Statens kartverk og kommunene, samt en rekke andre offentlige aktører som har et spesielt ansvar for geodata. Eksempler på slike aktører er Statens vegvesen, Telenor, Forsvaret og energiverkene.

---

[4] *Policy options for geographic information from Ordnance Survey*. Consultation. UK Department for Communities and Local Government, 2009

[5] *Re-use of Public Sector Information – Review of Directive 2003/98/EC*

Miljøverndepartementet legger til grunn at Norge digitalt har opphavsrettslig vern, og viser da til åndsverklovens databasevern. Det er dette som er bakgrunnen for at Norge digitalt kan kreve betalt ut fra hvilken nytte kundene har av dataene. Denne prisingsmodellen har vært kritisert for å fungere dårlig i forhold til tjenester på internett, som gjerne er gratis for brukeren. Med en modell som utløser en kostnad pr. visning, vil det være vanskelig for et nettsted å forutsi hvilke kostnader de vil få knyttet til visning av kart. De fleste nettsteder har derfor valgt å benytte kart fra andre leverandører, slik som *Google maps* eller *Open streetmap*. Disse holder ikke like høy kvalitet som kartene fra Statens kartverk og medfører at norske forbrukere får dårligere tjenester enn nødvendig.

Som en prøveordning åpnet Statens kartverk i desember 2009 sine karttjenester for utviklere som ønsker å lage applikasjoner basert på kart. Tjenestene som utvikles skal være gratis for sluttbrukerne. Kartverket har tidligere gjort kartene sin tilgjengelige for sluttbrukere gjennom tjenesten *Norgesglasset.no*.

2.3.2 Statistiske data

Statistisk sentralbyrå, SSB, forvalter store mengder statistikk knyttet til de fleste områder i samfunnet. Det er for eksempel SSB som mottar data som rapporteres fra kommuner og fylkeskommuner gjennom Kostra (Kommune Stat Rapportering).

Da *Fakta først-prosjektet* gjennomførte sin undersøkelse i 2009 snakket de med en rekke avdelinger i Hordaland fylkeskommune. Flere av avdelingene oppfattet det ikke som nødvendig å tilgjengeliggjøre egne data, fordi dataene ble tilgjengeliggjort i andre kanaler gjennom rapportering til SSB (Kostra) og Utdanningsdirektoratet.

Kostra-data inneholder blant annet statistikk om økonomi, skoler, helse, kultur, miljø, sosiale tjenester, boliger, tekniske tjenester og samferdsel. Tallene belyser prioriteringer, dekningsgrader og produktivitet

Testresultater for skoler rapporteres til Utdanningsdirektoratet og publiseres på *Skoleporten*[6], oftest på aggregert nivå slik at det er vanskelig å bearbeide dataene videre.

Mange av dataene SSB forvalter inneholder sensitive personopplysninger. Slike datasett bør anonymiseres før de tilgjengeliggjøres. SSB har også mulighet til å lage uttrekk hvor ulike datasett kobles sammen. SSB har tilgang til de personidentifiserbare dataene og kan gjøre koblingene, mens mottaker av datasettet får anonymiserte, statistiske data. For eksempel ønsker Senter for IKT i utdanningen i Tromsø å få laget en kobling mellom (anonymiserte) data fra skatteetaten og arbeidstakerregisteret for å kunne vise forventet inntekt for ulike yrkesgrupper. Slike koblede datasett bør i framtiden også gjøres tilgjengelig for andre etter at de er generert.

2.3.3 Forskningsdata

Et viktig argument i forhold til publisering av data er gjerne knyttet til at innsamlingen av dataene er finansiert av fellesskapet, og at dataene derfor bør komme fellesskapet til gode. En stor del av den forskningen som gjennomføres ved norske universiteter, høyskoler og forskningsinstitusjoner er finansiert gjennom offentlige forskningsprogrammer. Forskningsdata er likevel et av unntakene i gjenbruksdirektivet.

Tradisjonelt har forskningsresultater gjerne vært publisert i ulike vitenskapelige tidsskrifter som det har vært kostbart for allmennheten å få tilgang til. Disse tidsskriftene har gjerne lagt begrensninger på forskernes muligheter til å publisere sine resultater andre steder. De senere årene har imidlertid de institusjonene som finansierer forskningen lagt vekt på at resultatene skal være tilgjengelige for alle, gjerne i form av at forskningsprosjektet har et eget nettsted hvor alle publikasjoner offentliggjøres. Unntaket er der hvor det er forventet at forskningen skal kunne kommersialiseres, for eksempel i form av ny forretningsvirksomhet eller patenter.

[6] skoleporten.utdanningsdirektoratet.no

Til tross for at stadig flere publikasjoner fra forskningsprosjekter gjøres allment tilgjengelige, er underlagsdataene for forskningen i liten grad tilgjengelige. Det kan være rimelig å stille spørsmål ved denne praksisen. Data fra offentlig finansierte forskningsprosjekter – i den grad de ikke inneholder personsensitiv informasjon – bør også gjøres tilgjengelig for allmennheten. Dette vil kunne ha to positive effekter:

- Andre forskere kan bygge videre på eksisterende datasett, slik at man unngår dobbeltarbeid.
- Dersom det stilles spørsmål ved resultatet av forskningen, har kritikerne tilgang til grunnlagsdataene, og kan sjekke om de underbygger resultatet (etterprøvbarhet).

2.3.4 Kulturdata

Denne type data kan være film og teater produsert og finansiert av det offentlige, TV-program, materiale fra Riksarkivet eller Nasjonalbiblioteket, kartoteksystemer etc. Dette er data som ikke er omfattet av gjenbruksdirektivet på grunn av den omfattende problematikken knyttet til eierrettigheter til materialet.

NRK har de siste årene tilgjengeliggjort en stor del av sitt arkivmateriale for visning på nett, og også gjort forsøk med å gjøre enkelte programmer tilgjengelige for nedlasting i høy kvalitet. De har også eksperimentert med å oppfordre til gjenbruk av materiale der det ikke foreligger rettighetsproblematikk, for eksempel er programmet *Bergensbanen – minutt for minutt* gjort tilgjenglig i HD med en Creative Commons lisens som tillater både kommersiell- og ikke-kommersiell bruk av materialet, så lenge NRK krediteres.

I *Innovasjonsmeldingen* fra 2009 sier regjeringen at det skal legges fram en egen stortingsmelding om digital tilgjengeliggjøring av den delen av kulturarven som er bevart i arkiv-, bibliotek- og museumssektoren.

2.3.5 Operasjonelle data

Med operasjonelle data mener vi data som offentlige etater får tilgang til som en del av sitt daglige virke. Dette kan være data som passeringer gjennom bomsnittet i Oslo, trafikkinformasjon og lignende.

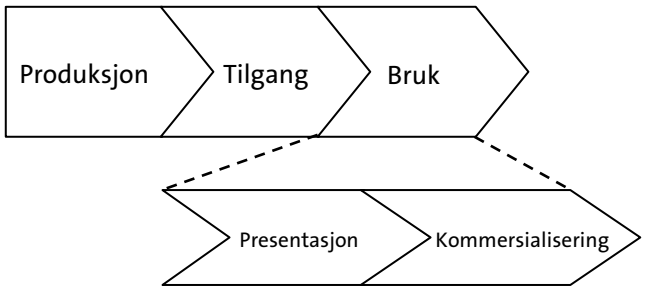
Andre eksempler kan være logger over søkeord på offentlige nettsider, besøkstall på ulike offentlige nettsider, statistikk over innsynsforespørsler og lignende. Denne type data bør i størst mulig grad gjøres løpende tilgjengelig.

Kapittel 3 | Forretningsmodeller for offentlige data

Til tross for at det er vanskelig å kvantifisere, synes det å være bred enighet om at offentlige data har stor verdi. I hvilken grad skal det offentlige ta ut denne verdien direkte? Skal de offentlige være en formidler av rådata, eller er det en offentlig oppgave å tilrettelegge dataene slik at vanlige brukere kan ha glede av dem?

3.1 Det offentliges plass i verdikjeden

EU-kommisjonen gjorde en evaluering av innføringen av gjenbruksdirektivet i 2009. Her fremgår det at mye av potensialet for gjenbruk ikke blir realisert fordi offentlige myndigheter har større fokus på kortsiktig inndekning av egne utgifter enn på langsiktige gunstige effekter på økonomien i samfunnet.



Figur 1: Verdikjedemodell for innovativ bruk av offentlige data[7]

Figuren viser en overordnet framstilling av verdikjeden for offentlige data. Selv om man ser bort fra problematikken med prising av data, kan man tenke seg at det offentlige opptrer flere steder i denne kjeden. Myndighetene står for produksjon i form av innsamling eller generering av dataene. Skal myndighetene utover dette kun være en leverandør av rådata, eller skal de også stå for presentasjonen i form av sluttbrukerløsninger, slik som *yr.no?*

Det er et uttrykt ønske, både fra dagens regjering og EU, at det offentlige skal tilgjengeliggjøre mest mulig rådata. Dette burde være uproblematisk så lenge det ikke truer personvernet eller rikets sikkerhet. Det synes imidlertid ofte som om offentlige myndigheter – inspirert av suksessen til yr.no – er mer opptatt av å generere positiv oppmerksomhet rundt sin egen virksomhet, og å generere trafikk til egne nettsider. Dette kan medføre et problem for portaler som ønsker å sammenstille mest mulig offentlig informasjon (eller offentlige tjenester, slik som *minside.no*), og det kan bidra til at det blir vanskeligere for sluttbruker å finne fram til de data eller tjenester som er tilgjengelige.

Et annet aspekt ved dette er at de ulike virksomhetene gjerne ønsker å kontrollere presentasjonen av sine data selv. Dette vanskeliggjør bruk av datasettene i sammenstillinger og analyser, noe som kan være ønskelig for å styrke informasjons- og ytringsfriheten og allmennhetens kontroll med det offentlige.

Det kan også være et problem at det offentlige tilbyr tjenester i konkurranse med private. Igjen kan yr.no brukes som eksempel, idet det her tilbys en gratistjeneste som det kommersielle foretaket Storm kunne ønsket å ta betalt for. Samtidig er det viktig å fremheve at i dette tilfellet tilbys dataene tjenesten baserer seg på til alle aktører i markedet på like vilkår, og yr.no er ikke en tjeneste som genererer inntekter i form av sluttbrukerbetaling eller reklame.

Det ville være problematisk dersom en offentlig tjeneste, finansiert over statsbudsjettet, tar betalt for sine tjenester, men holder så lave priser at den utkonkurrerer de andre i markedet. Det er derfor viktig at de dataene offentlige aktører baserer eventuelle sluttbrukerløsninger på også er tilgjengelig for andre på samme vilkår (det vil si til

[7] *Innovativ udnyttelse af offentlige data*. Rapport til IT-og Telestyrelsen og Forsknings- og Innovationsstyrelsen, 2009

nullkost). Da har kommersielle aktører mulighet til å konkurrere gjennom å tilby bedre og mer innovative løsninger.

Som hovedprinsipp bør likevel det offentlige være leverandør av data på et format som er egnet for gjenbruk, og overlate utvikling av tjenester til markedet. Dersom det offentlige likevel mener at det er på sin plass å tilby tjenester basert på egne data, bør ikke den offentlige tjenesten ha anledning til å kreve betaling fra sluttbruker, og tjenesten må heller ikke være reklamefinansiert. Dette er viktig for å sikre at markedet for verdiskaping i privat sektor ikke ødelegges, og man oppnår det motsatte av det som er intensjonen med gjenbruksdirektivet. Det er også et viktig prinsipp at offentlige tjenester er gratis, og det å bli eksponert for reklame er en form for "betaling" fra sluttbruke som det offentlige bør unngå. NRK har blant annet tatt konsekvensen av dette, og besluttet å fase ut reklame fra sine nettsider.

I EU-kommisjonens evaluering av innføringen av gjenbruksdirektivet vises det til at flere land som har endret prisstrategi på offentlige data har opplevd både økning i antall brukere, og økt variasjon i bruksmåter for offentlige data.

3.2 Pris på data skal være uavhengig av bruk

Offentlighetsloven, så vel som gjenbruksdirektivet, slår fast at det ikke er anledning til å forskjellsbehandle sammenlignbare tilfeller ved behandling av innsynsbegjæringer. I direktivet presiseres det imidlertid at dette ikke betyr at man ikke kan ha ulik pris for kommersiell og ikke-kommersiell bruk. Denne presiseringen er imidlertid ikke tatt inn i den norske offentlighetsloven.

I *NOU 2003:30 Offentlighetsloven* sier offentlighetsutvalget at det anser det som uaktuelt å la en betalingsplikt avhenge av om den som krever innsyn etter offentlighetsloven gjør det ut fra kommersielle motiver eller ikke. En slik regel vil blant annet medføre at forvaltningen må vurdere motivene til den som krever innsyn. Dette vil være både prinsipielt og praktisk svært uheldig, i følge utvalget. Offentlighetsutvalget hadde nok først og fremst tradisjonelt saksinnsyn i tankene, men problemstillingen er minst like aktuell i forhold til tilgjengeliggjøring av datasett.

Til tross for at datasett kan gjøres tilgjengelige med klart definerte modeller, er det ikke alltid like åpenbart hvilke regler som gjelder. Dette så vi tydelig da Statens Kartverk annonserte at de ville gjøre sine kart tilgjengelige for «personlig bruk». For hva er «personlig bruk»? Journalister og bloggere diskuterte friskt om dette for eksempel betød at det ikke ville være lov å bruke kartene på nettsider som var finansiert med reklame.

Sanntidsdata fra Trafikanten

Trafikanten er servicesenteret for kollektivtrafikk i Oslo. De siste årene har de fleste trikker, busser og T-baner i Oslo fått utstyr som angir hvor de befinner seg til enhver tid. På den måten kan de reisende få informasjon om hvor lenge det er til deres transportmiddel kommer til holdeplassen. Etter hvert som stadig flere har mobiltelefoner med GPS, slik at også den reisendes lokasjon er kjent, har det dukket opp et marked for mobilapplikasjoner som kobler sammen kunde og transportmiddel.

Trafikanten tilbyr et sett med webtjenester og programmeringsgrensesnitt (API) for utviklere, men disse er kun for ikke-kommersiell bruk. Dette betyr at det kun er utviklere som ønsker å lage gratistjenester som kan benytte seg av disse.[8]

For mange aktører i offentlig sektor synes det å være vanskelig å akseptere at noen kan tjene penger på en "råvare" de får gratis. Det koster for eksempel ikke Trafikanten mer å gi bort dataene dersom formålet er kommersielt. Begrunnelsen for å skille mellom de som ønsker å lage noe som er gratis, og de som ønsker å tjene penger, er kun at det oppleves som urettferdig om ikke Trafikanten får noe tilbake dersom noen andre tar betalt for deres data.

[8] Labs.trafikanten.no

Det er interessant at mange offentlige aktører synes å ha en motvilje mot at andre skal kunne tjene penger på "deres" data. Det er jo nettopp dette som er motivasjonen bak gjenbruksdirektivet: Gjennom å frigi offentlige data skal man bidra til innovasjon og verdiskapning, som igjen genererer inntekter til fellesskapet gjennom skatter og avgifter. I sin evaluering av innføringen av direktivet peker EU-kommisjonen nettopp på uvilje mot å la private næringsdrivende bruke offentlige data som en av faktorene som gjør det vanskelig å ta ut det fulle økonomiske potensialet i gjenbruk av offentlige data.

3.3 Databasevern og uforutsigbare betalingsmodeller

En del data det offentlige tilbyr i dag er ikke rådata, men data som tilrettelegges basert på etterspørsel fra næringslivet. Deler av dataene produseres også for å dekke dette behovet. Dette ser vi spesielt for geodata. Slike data, som det offentlige bruker ressurser på å samle inn, vedlikeholde og tilrettelegge for bruk har en spesiell status i forhold til andre offentlige data – det regnes som åndsverk, og er beskyttet av *databasevernet* som er omtalt i åndsverkslovens §43:

> *Den som frembringer et formular, en katalog, en tabell, et program, en database eller lignende arbeid som sammenstiller et større antall opplysninger, eller som er resultat av en vesentlig investering, har enerett til å råde over hele eller vesentlige deler av arbeidets innhold ved å fremstille eksemplar av det og ved å gjøre det tilgjengelig for allmennheten.*

Databasevernet er interessant i forhold til betalingsmodeller for offentlige data, fordi offentlighetsloven sier at data skal utleveres gratis med mindre det offentlige organet har en hjemmel i forskrift til å ta betalt. Betalingen skal i så fall gjenspeile de kostnader organet har hatt med å frembringe dataene, og i noen situasjonsbeskrevne tilfeller kan organet også beregne en rimelig avkastning av investeringene. Et særlig spørsmål er om organet kan ta betalt ut fra hvilken nytte kunden har av de kjøpte dataene.

Den typiske situasjon er at kunden kjøper elektroniske data fra et offentlig organ, og deretter lar disse dataene inngå i en nettbasert forretningsmodell, der inntjeningen er knyttet til antallet treff på nettstedet. Skal det da være lov å knytte betalingen opp mot antall treff, og på denne måten la suksess/manglende suksess være en parameter med hensyn til hvilken pris virksomheten må betale for dataene?

Dette spørsmålet er sentralt, fordi en slik betalingsmodell lett kan oppfattes å være på kant med retten til fri gjenbruk av data man har kjøpt. Offentlighetsloven gir anledning til fri gjenbruk av informasjon som er gjort tilgjengelig. Dette er bare begrenset i forhold til annen lovgivning (som for eksempel personopplysningsloven) eller dersom en tredjepart har rettigheter til informasjonen.

Justisdepartementet har i Kgl. res. av 2008-10-17 gitt uttrykk for at det bare er anledning til å knytte betalingen opp mot kundens nytte dersom den offentlige virksomheten har immaterielle rettigheter til informasjonen/dataene. Justisdepartementet sier ikke noe nærmere om hva slags immaterielle rettigheter som kan begrunne at prisen avhenger av bruken, men det er dette resonnementet som er bakgrunnen for at for eksempel kartdata er priset etter en slik modell.

Matrikkelinformasjon
Matrikkelinformasjon er tekniske opplysninger om eiendommer og bygninger. Informasjonen registreres når bygningen er ny og vedlikeholdes deretter ikke. Miljøverndepartementet lagt til grunn at det offentlige kan la betalingen for denne type informasjon være avhengig av hvordan opplysningene skal brukes og eventuelt videre-brukes. Dette begrunnes i databasevernet i åndsverksloven.

Prismodellen for matrikkelinformasjon synes å avvike fra Justisdepartementets beskrivelse av hvordan regelverket skal praktiseres, ettersom det er vanskelig å se at det finnes noen immaterielle rettigheter av betydning for denne typen infor-

masjon. Dersom myndighetene kan bruke databasevernet som argument for å prise matrikkelinformasjon i forhold til bruk, betyr dette i praksis at alle offentlige data som er lagret i databaser er omfattet av databasevernet. Dette kan umulig være i overensstemmelse med intensjonen verken i offentlighetsloven eller gjenbruksdirektivet.

3.4 Offentlige data bør være gratis

I utgangspunktet bør alle offentlige data som det ikke er gode grunner for å unnta offentlighet gjøres tilgjengelig gratis. Som tilbyder av data til fellesskapet bør ikke det offentlige stille betingelser knyttet til bruken og heller ikke skille mellom kommersiell og ikke-kommersiell bruk. Dette vil gi en enkel og ryddig forretningsmodell der det offentlige i hovedsak er tilbyder av (gratis) data, uten den administrasjonen som vil være nødvendig dersom man skal vurdere formålet med dataene, hvilken prismodell som gjelder i hvilket tilfelle etc. I stedet for å bruke ressurser på å drive forretning, kan det offentlige fokusere på sin kjernevirksomhet.

I de tilfellene hvor dataene som tilbys har mer preg av en *tjeneste*, for eksempel at dataene samles inn spesielt for å kunne dekke et behov i markedet, og ikke fordi det offentlige selv har bruk for dem, bør det være anledning til å ta betalt for dataene. Dette er tilfelle for en del av de tjenestene som er knyttet til geodata i dag. For øvrig bør det selvsagt være anledning til å ta betalt for fysiske tjenester som har en reell kostnad, slik som oppmåling av eiendom eller attesterte utskrifter fra offentlige registre.

Dersom det offentlige ikke ønsker å gjøre data hvor man i dag har god inntjening gratis, bør man som et minimum velge enkle modeller for kjøp av dataene, for eksempel etter modell fra Østerrike (se kapittel 1.4.2). For små oppstartsbedrifter er det viktig at det er lave og forutsigbare inngangskostnader. Det bør i så fall være enkle betalingsmodeller som ikke krever ekstra administrasjon, og det bør ikke være betingelser knyttet til bruken.

Erfaringene fra Danmark, hvor de etablerte en portal for offentlige data i 2009, er at de som ønsker å gjenbruke data ikke først og fremst er opptatt av om dataene er gratis, men av en god, fleksibel og forutsigbar forretningsmodell. Aktørene der ønsker seg mer modulariserte modeller, slik at man ikke behøver å betale for data man ikke trenger.

For de fleste offentlige etater vil tilgjengeliggjøring av data medføre små økte kostnader, og i enkelte tilfeller er det sannsynlig at det også kan føre til besparelser. Dette gjelder spesielt for virksomheter som får mange henvendelser om innsyn etter offentlighetsloven. Dersom informasjon og data er åpent tilgjengelig, vil man slippe mye av den kostbare saksbehandlingen som følger med behandling av innsynsbegjæringer.

Betalingsinnkreving er heller ikke gratis. Den kostnaden som følger med fakturering og oppfølging av en betalingsmodell vil bortfalle dersom man ikke lenger tar betalt for dataene. Dette vil kunne bidra til å redusere administrative kostnader.

Det oppstår en spesiell utfordring når det ikke er den delen av virksomheten som har utgiftene (med å samle inn eller produsere data) som får inntektene (fra skatter og avgifter) knyttet til å frigi dataene. I forhold til kartdata har kommunene i dag anledning til å ta betalt for kart- og eiendomsinformasjon. På denne måten sikres de tilbakeføring av midler til ytterligere produksjon og videreutvikling av dataene. En bekymring fra kommunenes ståsted, dersom det blir åpnet for gratis kart- og eiendomsdata, er at intensjonen om økt aktivitet i næringslivet som følge av gratis data ikke genererer midler tilbake til kommunene. Eventuell økt aktivitet vil gi inntekter til staten i form av skatt og avgifter og mindre eller ingenting til kommunene.

For slike virksomheter vil det være nødvendig med økte overføringer fra sentrale myndigheter. På sikt er likevel tanken at disse utgiftene vil bli dekket inn gjennom økt samfunnsnytte og økte inntekter til staten gjennom skatter og avgifter fra nye virksomheter som tar i bruk dataene.

Kapittel 4 | En norsk portal for offentlige data

Det er tilsynelatende bred enighet – både i offentligheten og i politiske miljøer – om at det vil være et gode for samfunnet å tilgjengeliggjøre flest mulig offentlige datasett. Gjenbruksdirektivet peker på at dersom offentlige data skal være mest mulig tilgjengelige, må det også være mulig å finne fram til de dokumentene og datasettene som finnes. Direktivet pålegger de enkelte land å legge til rette for dette, for eksempel gjennom nettressurser.

Fordi offentlighetsloven først og fremst er uformet med tanke på tradisjonelt dokumentinnsyn, kan det være vanskelig å forstå hvilke datasett som omfattes av loven, og hvordan de skal tilgjengeliggjøres. Dette gjelder både for potensielle leverandører av data, og i forhold til brukerne. I tillegg er det et problem at man ikke vet hvilke datasett som faktisk finnes.

I en undersøkelse foretatt av Universitetet i Bergen i 2009, framgår det at et mindretall (32%) av de offentlige virksomhetene som ble undersøkt opplyser om egne datakilder på forsiden på sitt nettsted. 61% opplyser om egne datakilder et eller annet sted på sidene sine[9]. Det er derfor viktig at myndighetene får på plass en løsning for å synliggjøre hvilke datasett som er tilgjengelige.

I EU-kommisjonens evaluering av innføringen av direktivet pekes det på flere hindringer for å få tatt ut det fulle potensialet i offentlige data. To av disse er spesielt relevante i denne sammenheng:

- Markedet har manglende informasjon om offentlige datasett
- Offentlige organer har manglende forståelse for gjenbrukspotensialet i egne data

Vi foreslår at det så snart som mulig etableres en portal som gir oversikt over offentlige datakilder i Norge. Dette vil kunne avhjelpe den første av disse hindringene.

4.1 Hvordan bør portalen se ut?

Det bør etableres en portal som tilbyr en oversikt over hvilke offentlige datakilder som fins, med tilstrekkelig informasjon til at brukeren kan få en forståelse for hvordan dataene kan brukes. Slik informasjon vil være:

- Beskrivelse av dataene
- Hvilket format dataene finnes på
- Beskrivelse av eventuelt programmeringsgrensesnitt (API) som kan brukes for å få tilgang til dataene
- Tidsangivelse for et datasett, evt. om det dreier seg om sanntidsdata
- Hvilken kvalitet dataene har

4.2 Ulike formateringsnivåer

Vi kan skille mellom ulike formateringsnivåer på dataene:

Rådata er data på det formatet det brukes hos det aktuelle organet. Begrepet brukes på litt ulike måter i ulike miljøer, men når vi her snakker om rådata mener vi også data som for eksempel er blitt anonymisert, så lenge dataene foreligger på et maskinlesbart format. Dersom det er snakk om samme datasett, men det tilbys på flere ulike formater, vil vi plassere alle versjonene i denne kategorien. Det samme gjelder for et API som gir direkte tilgang til dataene.

Presentasjonsdata er data som er tilrettelagt slik at de er enklere å ta i bruk for sluttbruker, for eksempel som et bilde av et kartutsnitt. Også presentasjonsdata kan tilbys gjennom et API, slik

[9] *Fakta først. Viderebruk av datakilder i offentlig sektor: potensial og hindringer*, Prosjektrapport, Institutt for informasjons- og medievitenskap, UiB

som Statens kartverk gjør med sin løsning for utvikling av ikke-kommersielle tjenester.

Tilrettelagt for sluttbruker. Dette er når en offentlig virksomhet selv utvikler applikasjoner basert på egne data beregnet for sluttbrukere, slik som yr.no og Trafikanten. Ulike grensesnitt for å begrense tilgangen der det på grunn av hensyn til personvernet ikke er forsvarlig å publisere hele datasettet vil også falle inn i denne kategorien (se kapittel 2.1.2).

Vi mener at det offentlige minimum må gjøre sine rådata tilgjengelige. Det et også ønskelig at dataene gjøres tilgjengelige på formater som forenkler gjenbruk. For data hvor dette er relevant bør det finnes et godt dokumentert API. Det må tilbys et API for alle typer sanntidsdata.

I utgangspunktet bør ikke det offentlige konkurrere med næringslivet og frivillige krefter om å lage best mulig sluttbrukerapplikasjoner, med unntak av de tilfellene vi har beskrevet i kapittel 2.1.2

Det er imidlertid viktig å ikke vente til alt er perfekt eller alle problemstillinger er avklart. En portal bør etableres og befolkes med kjente datasett så snart det er praktisk mulig. Deretter kan det gjøres løpende forbedringer, både i forhold til antall datasett og hvilke formater og API-er som finnes.

4.3 Hvem skal forvalte portalen?

Det synes mest hensiktsmessig at portalen settes opp som en datakildekatalog, og at dataene forvaltes hos den enkelte virksomhet. Fra Danmark har man erfart at det er viktig at de ulike dataeierne også får et forhold til etterspørselen etter sine data.

Det er viktig at de ulike offentlige virksomhetene har et fast kontaktpunkt hvor de kan få råd og veiledning i forhold til å gjøre datasett tilgjengelige, for eksempel knyttet til hva som skal unntas fra offentlighet. Det er også viktig at brukerne har et kontaktpunkt inn mot det offentlige hvor de kan henvende seg dersom de er usikre på hvem som forvalter de dataene de trenger.

Det bør vurderes hvem som skal ha hovedansvaret for en slik datakildekatalog. Direktoratet for forvaltning og IKT (Difi) kan være en aktuell kandidat. Statistisk sentralbyrå forvalter i dag store mengder data som rapporteres inn fra ulike deler av forvaltningen. Byrået har også lang erfaring knyttet til anonymisering av datasett og håndtering av forespørsler om innsyn i data. Også Brønnøysundregistrene har lang erfaring i å forvalte offentlige data.

Den etaten som skal forvalte datakildekatalogen bør ha et mandat både til å aktivt kartlegge hvilke datakilder som finnes, men også til å jobbe for at flest mulig datakilder gjøres tilgjengelige. Det er flere virkemidler som kan brukes:

Veiledning og insentiver
Først og fremst bør det legges til rette for at dataeierne kan gjøre sine data tilgjengelige og legge inn referanse til datasettet/API-et i katalogen. Dette kan gjøres gjennom å aktivt oppmuntre alle offentlige myndigheter til å vurdere sine datasett for publisering. Der det er aktuelt bør API-er og rutiner for tilgjengeliggjøring innarbeides i kravspesifikasjonene når nye systemer skal kjøpes inn eller utvikles.

Det bør utarbeides en håndbok med retningslinjer for hvordan virksomheter skal gå fram når det skal gjøre data tilgjengelige. I hvilken grad data gjøres tilgjengelig bør være et av kriteriene i Norge.nos vurdering av offentlige nettsteder.

Det er naturlig at det overordnede ansvaret er forankret hos Fornyings-, administrasjons- og kirkedepartementet. Erfaringene fra Danmark og Storbritannia viser imidlertid at gjennomslagskraften blir enda større dersom prosjektet i tillegg er forankret på høyeste nivå i Regjeringen. I Storbritannia har Gordon Brown selv vært frontfigur for lanseringen av *data.gov.uk*.

Aktiv oppfølging
Det vil være en uoverkommelig oppgave dersom en etat skulle ha ansvaret for å finne fram til alle mulig datasett i landet. Gjennom tilrettelegging som

nevnt over, bør de ulike dataeierne selv innrapportere sine datasett. Forvalteren av portalen bør imidlertid ha gode analyseverktøy som kan danne grunnlag for oppfølging mot alle de som ikke innrapporterer sine data, til tross for gode veiledere. Etaten bør følge opp dersom det er aktører som ikke har innrapportert *noen* datasett. Dersom en eller flere kommuner rapporterer inn en type datasett som det er rimelig å anta at alle kommuner besitter, bør dette følges opp mot de kommunene som ikke har innrapportert tilsvarende datasett.

Videreformidling av publikumshenvendelser
Dersom noen som ønsker å gjenbruke data ikke finner de dataene de trenger i portalen, bør de kunne henvende seg til etaten med en forespørsel om å få offentliggjort det aktuelle datasettet. Etaten bør da behandle dette som en vanlig innsynsbegjæring, og videreformidle denne til til dataeier. Dataeier må da behandle denne og ta stilling til om dataene kan gjøres tilgjengelige. Dersom en sluttbruker etter en slik henvendelse får innsyn i et datasett som tidligere ikke har vært, bør hovedregelen være at dette datasettet deretter gjøres tilgjengelig for alle gjennom datakildekatalogen.

4.3.1 Finansiering
Finansieringen bør tilrettelegges slik at dataeierne i minst mulig grad får ekstrautgifter – og at eventuelle gevinster ved tilrettelegging av data (for eksempel ved at man oppfyller dokumentasjonskrav) tilfaller dataeieren. Dette for å rydde av veien et av de viktigste hindrene for dataåpenhet.

4.4 Hvor skal vi begynne?

Det finnes en rekke offentlige datasett som egner seg for offentliggjøring. Så snart en portal for offentlige data er etablert, bør de etatene som har slike data begynne å gjøre dem tilgjengelige.

Det er imidlertid enkelte datasett det knytter seg mer interesse til enn andre, og vi vil anbefale at den som forvalter portalen arbeider aktivt for å få på plass disse så raskt som mulig. Den opplevde nytten av portalen vil være større dersom den tidlig inneholder datasett som oppleves som relevante for brukerne. Dette vil også kunne bidra til at man ser raske resultater knyttet til etableringen av portalen. Vi foreslår at det fokuseres spesielt på følgende datasett:

Samferdselsdata
Dette kan være data knyttet til rutetider for lokaltransport og fjerntransport, sanntidsinformasjon for lokaltransport, sanntids trafikkinformasjon fra Statens vegvesen med mer. Dette er data som har mange praktiske anvendelser, og som det er stor interesse for i markedet.

Kartdata
Alt blir bedre når man kan knytte det til lokasjon. Kartdata er nøkkelen til gode, brukervennlige tjenester.

Kvalitets- og effektivitetsdata for offentlig sektor
Hvordan er ventetiden ved mitt lokalsykehus? Hvor mange kneoperasjoner foretok de i fjor? Hvor fornøyde er foreldrene som har barn i "Hompetitten" barnehage? Hvilken kommune har lavest kommunale avgifter? Er det noen sammenheng mellom nivået på kommunale avgifter og hvor fornøyde brukerne er?

Offentlige kvalitets- og effektivitetsdata er interessante både for sluttbrukere og for en kritisk og undersøkende presse.

Litteratur

Innovativ udnyttelse af offentlige data. Rapport til IT-og Telestyrelsen og Forsknings- og Innovationsstyrelsen, 2009

Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information

Et nyskapende og bærekraftig Norge. St.meld. nr. 7 (2008-2009)

Exploiting government's information assets for the public good. Advisory Panel on Public Sector Information (APPSI), UK Ministry of Justice, 2009

Fakta først. Viderebruk av datakilder i offentlig sektor: potensial og hindringer. Prosjektrapport fra Institutt for informasjons- og medievitenskap ved UiO, 2010

For innholdsleverandørene til OEP (Offentlig elektronisk postjournal). Veiledning i reglene om arkivering og journalføring og om taushetsplikt fra Fornyings- og administrasjonsdepartementet, 2009

Fra bruk til gjenbruk. Tilrådinger fra interdepartmental arbeidsgruppe til Arbeids- og administrasjonsdepartementet, 2004

Iverksetting av lov 19. mai 2006 nr.16 om rett til innsyn i dokument i offentleg verksemd og forskrift til lova. Kongeleg resolusjon 17. oktober 2006

MEPSIR. Measuring European Public Sector Information Resourses. Dekkers, Polman, te Velde, de Vries for EU-kommisjonen, 2006

Models of Public Sector Information Provision via Trading Funds. Newbery, Bentley og Pollock, Cambridge University for UK Department for Business, Enterprise and Regulatory Reform (BERR) and HM Treasury, 2008

"Norge digitalt" – et felles fundament for verdiskapning. St.meld. nr. 30 (2002-2003)

Om gjennomføring av Inspire-direktivet. Høringsnotat fra Miljøverndepartementet, 2008

Om lov om rett til innsyn i dokument i offentleg verksemd (offentleglova). Ot.prp. nr. 102 (2004-2005)

Policy options for geographic information from Ordnance Survey. Consultation. UK Department for Communities and Local Government, 2009

Rettleiar til offentleglova. Justis- og politidepartementet, 2010 (endelig versjon)

Re-use of Public Sector Information – Review of Directive 2003/98/EC. Communication from the Commission to the European Parliament, the council, the European Economic and Social Committee and the Committee of the Regions, 2009

Utkast til forskrift om rett til innsyn i offentleg verksemd (offentlegforskrifta), Høringsnotat fra Lovavdelingen i Justis- og Politidepartementet, 2007