

REFERANSE:
14/7257

DATO:
17.02.2015

VERSJON:
3.0

ANSVARLIG:
Eszter Horvati

Kartlegging av åpne data

av relevans for byggesaksprosessen

Innhold

1	Innledning	3
1.1	Hva er åpne data?	3
1.2	Oppdragets mandat	3
2	Gjennomføring av oppdraget	4
3	Selvetjent byggesaksprosess	6
3.1	Garasjebyggeren Sjur v.1.0	6
3.2	Garasjebyggeren Sjur v.2.0	6
3.3	Bruken av åpne data	7
4	Oversikt over funnene	9
4.1	Formater	9
4.2	Tilgjengelighet	9
4.3	Bruk av lisenser	9
5	Variasjon i tilnæringsmåte	11
5.1	Problemer og hindringer	11
5.2	Kommunikasjon med kommunene	12
6	Konklusjon	14
	Vedlegg	15

1 Innledning

Bouvet har på oppdrag for Direktoratet for byggkvalitet (DiBK) gjennomført en kartlegging over hvilke data og datasett som relevante sektormyndigheter involvert i byggesaksprosessen forvalter. Målet var å avdekke hvilke grunnlagsdata som på kort og noe lenger sikt kan understøtte selvbetjeningsløsninger i byggesaksprosessen.

Innsamlingen gjaldt alle datasett av relevans, både åpent tilgjengelige og de som enten er tilgjengelige under visse begrensninger eller overhodet ikke er tilgjengelige for andre enn dataeier. Det var også viktig å kartlegge hvilke type ressurser det dreier seg om og hvorvidt det er relativt enkelt å få dataene over på en hensiktsmessig maskinlesbar form, der man i dag bare har proprietære eller «menneskelesbare» formater.

1.1 Hva er åpne data?

Direktoratet for byggkvalitet ønsker å bidra til at informasjonsflyt i forbindelse med byggesaksprosessen blir enhetlig og harmonisert. Det innebærer blant annet at man legger vekt på å øke gjenbruk og viderebruk av offentlig informasjon på tvers av myndigheter. Digital informasjon som er gjort tilgjengelig for viderebruk av alle, kalles for *åpne data*. Åpne datasett gir mulighet til å skape nye innovative tjenester. En rekke datasett fra offentlig sektor er allerede tilgjengelige i forskjellige formater på nettstedene data.norge.no og GeoNorge (www.kartverket.no/Geonorge). Det finnes likevel mye nyttig informasjon og mange registre som ikke er åpent tilgjengelige, eller ikke er i et format som enkelt kan leses av maskiner. Åpne data for bruk til automatisert byggesaksbehandling bør være av høy kvalitet, oppdatert og relevant. På den måten kan man sikre at borgerne blir likt behandlet, uansett hvor de bor.

1.2 Oppdragets mandat

Oppdragets mandat har vært å kartlegge, gjennom telefonintervjuer, hvilke datasett som er åpent tilgjengelige, og å fastslå status på andre datasett ut fra alle relevante datasett hos de 15 sektormyndighetene som kan være berørt i en byggesak.

Det var også et mål å gi en form for analyse og vurdering av innhentet informasjon i forhold til tre områder:

1. I hvilken grad andre datasett på en enkel måte kan gjøres tilgjengelige;
2. Identifisere og beskrive eierskap for datasettene;
3. Avhengigheter mellom datasettene.

For å sammenstille funnene, ble det utarbeidet en intervjuguide med en rekke spørsmål. Denne guiden finnes som egen fil vedlagt rapporten. Funnene fra undersøkelsen er presentert i denne rapporten og i en oversiktstabell, vedlagt som egen fil.

2 Gjennomføring av oppdraget

I alt omfattet kartleggingen 15 sektormyndigheter som er berørt av kommunens samordningsplikt etter plan- og bygningsloven § 21-15. Kartleggingens første fase var å identifisere de faginstansene som samler, produserer, digitaliserer og forvalter data for den aktuelle myndighet. Faktagrunnlaget som myndighetene baserer beslutningene på forvaltes i mange tilfeller av forskningsinstitusjoner utenfor selve myndighetene som fører tilsyn eller gir godkjenning. De siste årene har det vært en rekke omorganiseringer og sammenslåinger av etater og institusjoner i offentlig sektor. Dette har komplisert oppgaven med å identifisere kilden til faktagrunnlaget for sektormyndighetene. Det var derfor nødvendig å kartlegge også landskapet av dataforvaltere under de respektive myndighetene som kan være berørt i byggesaksprosessen. Kartleggingen omfatter ikke alle potensielle datasett hos alle relevante dataforvaltere, men gir en bred oversikt over situasjonen i dag.

Listen under viser en sammenstilling av titlene på de respektive sektormyndighetene (slik de er omtalt i plan- og bygningsloven) og deres dataforvaltere:

Myndighet	Dataforvalter
Vegmyndighet	Statens Vegvesen (SVV)
Forurensingsmyndighet	Miljødirektoratet (MDIR), Norsk institutt for luftforskning (NILU)
Friluftslivmyndighet	Miljødirektoratet, Norges vassdrags- og energidirektorat (NVE), Artsdatabanken
Jordlovmyndighet	Norsk institutt for skog og landskap, Landbruksdirektoratet (LDIR)
Reindriftsmyndighet	Norsk institutt for skog og landskap, Landbruksdirektoratet
Bergmyndighet	Norges geologiske undersøkelse (NGU), Direktoratet for mineralforvaltning (DMF)
Helsemyndighet	Folkehelseinstituttet (FHI), Mattilsynet, Statens strålevern (NRPA)
Brannvernmyndighet	Direktoratet for sikkerhet og beredskap (DSB)
Sivilforsvaret	Direktoratet for sikkerhet og beredskap
Havnemyndighet	Kystverket
Kulturminnemyndighet	Riksantikvaren, Norsk institutt for kulturminneforskning (NIKU)
Luftfartsmyndighet	Avinor
Jernbanemyndighet	Jernbaneverket (JBV)
Arbeidsmiljømyndighet	Arbeidstilsynet
Kirkelig myndighet	Kirkelig arbeidsgiver- og interesseorganisasjon (KA)

Det bør påpekes at mange av dataforvalterne samarbeider om datafangst og digital fremstilling av disse på tvers av flere fagområder. Store nasjonale databaser som Naturbase og NADAG (Nasjonal database for grunnundersøkelser) er eksempler på dette.

I denne kartleggingen har man ikke undersøkt hvor beslutningsmyndigheten i de enkelte områdene ligger. Den kan nemlig finnes både lokalt hos kommunen, regionalt hos fylkeskommunene og på nasjonalt nivå ved et direktorat. Eksempelvis har kommunene myndighet til å ta avgjørelser i

kulturminnesaker som berører SEFRAK-registrerte bygninger fra 1850 og nyere, fylkeskommunene fra mellom 1650 og 1850, og Riksantikvaren for automatisk fredede byggverk fra før 1650. Eksempelet viser at landskapet over myndighetsutøvelse for de enkelte områdene er svært omfattende.

Listen over dataforvaltere dannet grunnlaget for hvor mange intervjuer som måtte gjennomføres. Man forsøkte å komme fram til personer som kjenner dataene og datasettene godt, og har kunnskaper om hvordan de blir samlet inn og videre bearbeidet. Intervjuene ble i de aller fleste tilfellene gjennomført over telefon.

Oversikt over kontaktpersoner som har deltatt i kartleggingen:

Arbeidssted	Navn
SVV	Hilde Austlid
DSB	Karen Lie
DSB - Sivilforsvaret	Steinar Skaugvoll
Kystverket - Geodatatenesten	Simen Slotta
Riksantikvaren	Christer Vinje Gimse
Skog og landskap	Ingvild Nystuen
LDIR - Avd. for reindrift	Alf Bjørnar Eriksen
NGU	Inger-Lise Solberg
Avinor - Dataforvaltning	Margaret Giffen
JBV - Kartløsninger	Jon Haugland
JBV - Miljøseksjonen	Trygve Aasen
JBV - Eiendomsavdelingen	Torbjørn Immerstein
Arbeidstilsynet - Østfold og Akershus	Aleksander Ribe
Arbeidstilsynet - Styring og samordning	Heidi Blekkan
MDIR - Miljødataseksjonen	Ingunn M. Limstrand
Helsedirektoratet	Finn Martinsen
Mattilsynet	Jo Giske
FHI – Luft og støv	Per Schwarze
FHI – Inneklima	Johan Øvrevik
NRPA - Radon	Ingvild Finne
KA - Kirkebygg og kulturminneforvaltning	Kjersti Kambestad

Intervjuene ble gjennomført basert på en intervjuguide utarbeidet i samarbeid med DiBK. Intervjuguiden er vedlagt, se vedlegg 2.

Det har vært løpende kommunikasjon mellom DiBK og Bouvet om fremdrift og resultater i prosjektet. Omfanget av materialet var svært stort, så arbeidets fokus har vært å få avdekket flest mulig datakilder innenfor ressurs- og tidsrammen. Vi har altså lagt vekt på bredde heller enn dybde.

Kartleggingen ble gjennomført i perioden november-desember 2014. Rapporten er utarbeidet i januar 2015.

3 Selvbetjent byggesaksprosess

I direktoratets strategi ByggNett, er det et uttalt mål å få til et bedre samspill mellom kommunene, myndighetene, fagmiljøene og byggenæringen for å utvikle digitale tjenester som bidrar til forbedret byggkvalitet. ByggNetts overordnede visjon er *enkle, effektive og forutsigbare tjenester i hele byggverkets levetid*. Gjennom tilgjengeliggjøring, deling og sammenkobling av data på tvers av myndigheter og fagmiljøer, kan IKT-leverandører stimuleres til å utvikle tjenester som samler, knytter sammen og formidler relevant informasjon gjennom hele byggesaksprosessen. Man ønsker å avdekke om det foreligger muligheter for å realisere digitale selvbetjeningsløsninger for å lette og effektivisere byggesaksprosessen.

3.1 Garasjebyggeren Sjur v.1.0

For å belyse omfanget av mulighetene som deling og sammenkobling av åpne data gir, vil vi bruke et tenkt scenario utviklet av DiBK: «*Fremtidsvisjonen om garasjebyggeren Sjur*». Her er fire roller beskrevet:

- selvbyggeren Sjur Stokkvik, som vil bygge en dobbeltgarasje på tomten sin,
- naboene til Sjur, Kari og Arne Antonsen,
- byggesaksbehandleren ved Nedregård kommune, Knut Svingen, og
- arkivlederen ved Nedregård kommune, Karla Karlsen.

Felles for dem alle er at de har behov for å innhente bakgrunnsinformasjon for å kunne foreta vurderinger. For enkelhets skyld beskrives kun Sjurs behov i denne gjengivelsen.

Sjur sitter hjemme i sofaen og planlegger å bygge en dobbelgarasje på tomta si. Han blir oppmerksom på en digital tjeneste som han kan logge seg inn på med MinID. Her ser han straks en rekke fakta om eiendommen sin(1). Gjennom tjenesten kan han lage seg en 3D-skisse av garasjen med port, vinduer og dør. Tegningen plasserer han inn i et kart som viser tomta hans med både grensene til naboen og innkjørselen fra veien tydelig. Han får opp en boble som sier: «*Plassering av garasjen nærmere enn 1 meter fra nabogrense vil kreve ekstra nøye vurdering fra kommunens side og eventuelt brannisolering.*»(2) Han drar garasjen litt bort fra nabogrensen og nærmere veien. Da kommer en ny boble opp: «*Du har plassert bygningen nærmere veien enn Vegloven tillater. Tilstøtende vei er kommunal og byggegrensen er 15 meter målt fra midten av veien.*»(3) Til slutt kommer det opp spørsmål om det skal være profesjonelle arbeidsplasser i huset (4). Han krysser av for «nei».

I dette scenariet har tjenesten hentet inn data fra Statens Kartverk og Matrikkelen (1), brannvernmyndighet(2), vegmyndighet(3) og arbeidsmiljømyndighet(4) og vist den i konteksten der den blir relevant.

3.2 Garasjebyggeren Sjur v.2.0

La oss forestille oss en utvidelse av historien:

Sjur tenker at det kunne være fint å bygge garasjen som en forlengelse av hovedhuset. Han prøver ut dette på kartet ved å flytte modellen inntil husveggen. En melding kommer opp: «*Bygningen du planlegger å utvide er fra byggeår 1873 og ført på Byantikvarens liste over verneverdige kulturminner. Det må søkes om dispensasjon for å utføre endringer i bygningens struktur.*»(5) Samtidig dukker en annen melding opp: «*På området du ønsker å bygge er det registrert utbredelse av stær. Denne arten er oppført på Artsdatabankens rødliste som «nær truet». Fjerning av trær til fordel for byggetiltaket vil få konsekvenser for arten.*»(6) Sjur hever på øyebrynene. Dette hadde han virkelig ikke tenkt på! Han vurderer om det ikke hadde vært mer aktuelt å ha garasjen på den andre siden av tomta, den mot jernbanestasjonen. Han flytter bygget på tegningen. Nå dukker en ny melding opp: «*Du har plassert bygningen i et område nær jernbanen. Vær oppmerksom på at byggegrensen mot jernbanen er 30 m og at togtrafikk kan forårsake støypenger.*»(7) Sjur må smile. Den støyen han opplever fra togene er uvesentlig, dessuten har det ikke noe å si for garasjen, siden det ikke skal bo

noen der. Under hele prosessen har det blinket en liten varselampe ved garasjetegningen. Nå klikker han på den, og et kart med følgetekst kommer opp: «Byggingen kan berøre ledninger og kabler i grunnen. Ta en kikk på kartet over kjente kabler og rør i bakken, så du kan planlegge og utføre gravearbeid uten fare for liv, helse, materiell og eiendom. Du kan bestille en påvisning av hvor kablene nøyaktig ligger ved å klikke på [denne](#) lenken.»⁽⁸⁾ Sjur studerer kartet, og ser at det ikke er noen store ledninger der han har plassert garasjen nå, men de er ikke lenger unna enn at han velger å klikke på knappen «Send gravemelding til ditt lokale strømnetselskap samtidig med innsending av søknaden». Dermed kan han gå videre. Han blir riktig så imponert over informasjonen og servicen i denne tjenesten.¹

I denne versjonen har mer informasjon blitt konsultert, denne gangen fra kulturminnemyndighet(5), friluftslivsmyndighet(6), jernbanemyndighet/forurensingsmyndighet(7) og Norges vassdrags- og energidirektorat (NVE). De fleste byggesaker involverer kun et fåtall andre myndigheter, likevel sier dette bildet noe om kompleksiteten som ligger i informasjonsinnhenting i en byggesak. I eksempelet over har åpent tilgjengelige datasett fra forskjellige leverandører tilført direkte verdi til søkeren gjennom en digital selvbetjeningsløsning.

Videre i prosessen vil data og informasjon som understøtter saksbehandlerens arbeid komme til nytte. I tillegg til faktagrunnlag, kan man se for seg kontekstbasert informasjonsinnhenting som gjelder aktuelle lover og forskrifter, veiledninger til tolking av regelverket og eksempler på hvordan liknende saker har vært løst tidligere.

Generelt kan man si at data som tilflyter søkeren ikke bare vil forenkle innsending av søknaden, men også redusere antall mangelfullt utfylte søknader og dermed spare byggesaksbehandlerne for mye ekstra arbeid. Tiden brukt på faktainnsamling kan reduseres betraktelig ved å opprette mekanismer som henter «levende» data fra de åpne datakildene og retter dem inn mot de aktuelle problemstillingene i den enkelte sak.

3.3 Bruken av åpne data

Problemløsningsprosessen rundt byggesak stiller svært høye krav til tilgang og sammenstilling av informasjon. Å stille «diagnosen», dvs. å avgjøre hvilke kilder som skal konsulteres i hvilke tilfeller, er en krevende øvelse der tilgjengelighet av gjenbrukbare data fra mange kilder vil være av stor verdi.

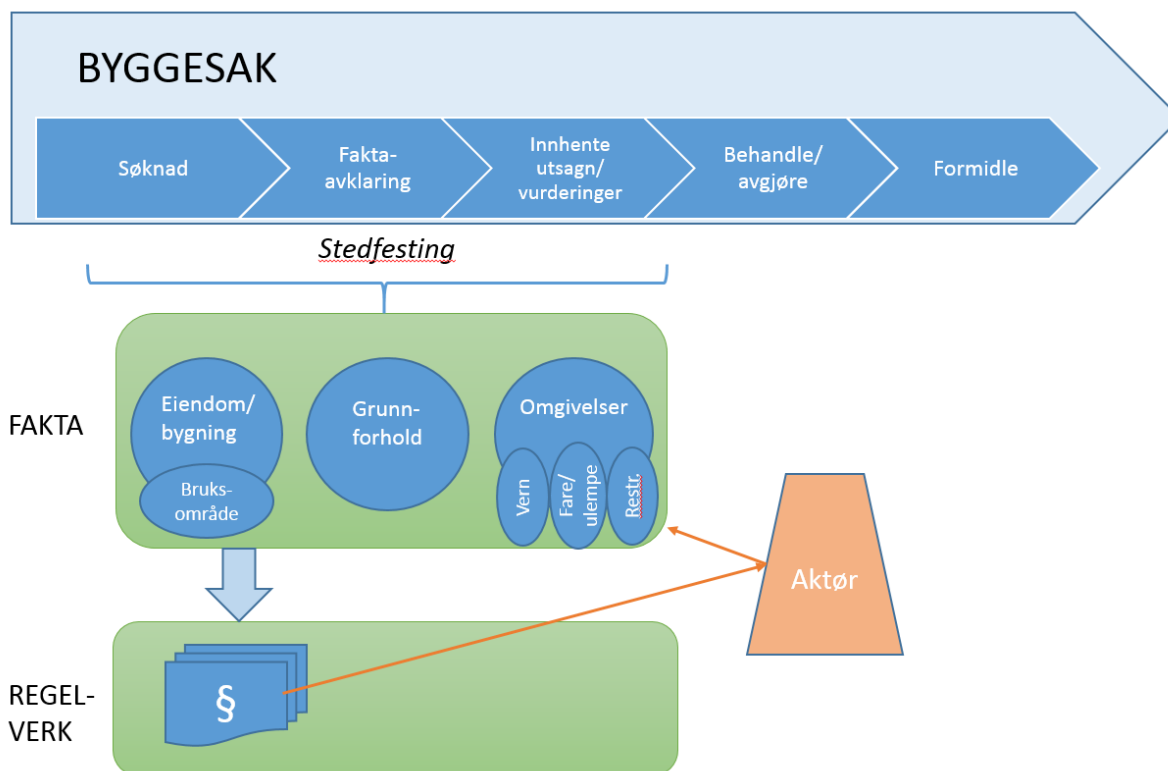
Ideelt sett skal disse dataene kunne sammenstilles på en meningsfull måte. Utfordringen med slike sammenstillinger er at data gjerne kommer i ulike formater og uten en eksplisitt «tolkningsnøkkel». Modeller som beskriver dataene og hva de handler om er viktige byggestener som letter bruken og sammenkoblingen av datasett fra ulike leverandører. SOSI (Samordnet Opplegg for Stedfestet Informasjon) er den største nasjonale standarden for geografisk informasjon, og et godt eksempel på en slik modell.

Det er tre klare fordeler med publisering av åpne data:

1. Man senker terskelen for bruk av dataene utenfor det formål de er skapt for. Så lenge datasettene er publisert i et maskinlesbart format, kan tredjepartsutviklere jobbe videre med dem. Dette motiverer til innovasjon. Dataformidlerne kan ikke forestille seg alle fremtidige applikasjoner og løsninger deres data kan være nyttige i.
2. Semantiske standarder sikrer riktig tolkning av data ved gjenbruk. Ved bruk av stabile, entydige identifikatorer kan man unngå begrepsforvirring og feil bruk av data.
3. Økt bruk vil øke datakvaliteten på kilden. Dataformidlerne bør betrakte eksponering som et gode, en unik mulighet til å få korrigert datakvaliteten.

Modellen under gir en illustrasjon på hvordan en selvbetjent byggesaksprosess kan dra nytte av åpne data:

¹ Eksempelet er inspirert av Lyse Elnetts veiledning til kunder som vil starte byggearbeider:



De tre første (forenklede) fasene i byggesaksprosessen konsentrerer seg om innhenting av fakta basert på stedfesting. Statens kartverk spiller gjennom Matrikkelen og Det offentlige kartgrunnlaget en sentral rolle i formidlingen av data relevant for byggesak. I Matrikkelen ligger essensiell informasjon om eiendommen, mens Det offentlige kartgrunnlaget samler geografiske data til bruk i plan- og byggesaker. Stedfesting er det viktigste koblingselement i datagrunnlaget. Kildene for å fatte et vedtak kan være faktagrunnlag som registre og databaser, eller lover, regler, forskrifter og veiledninger til disse.

Faktadelen kan man dele i tre:

- Informasjon om eiendommen (nybygg eller endring på eksisterende bygg), f.eks. byggeår, eierskap, nøyaktige mål, bruksområde.
- Data om grunnforhold, f.eks. veg- og jernbanenettet, ledninger, naturinfo, dyrkbar jord.
- Informasjon om omgivelsene igjen delt i temaene *vern* (f.eks. naturvernområder, rødlistede arter, verdifulle kulturlandskap), *fare/ulempe* (f.eks. skredinformasjon, radon, støysoner), og *restriksjoner* (f.eks. byggerestriksjoner i nærheten av lufthavn, strandsonen, hensynssoner ved eksplosivanlegg).

Aktør i denne figuren kan dekke ulike roller:

- Dataeier/-forvalter – den som avleverer fakta på forespørsel.
- Regelverksforvalter / mulig høringsinstans – de som ut fra sitt ansvar og/eller et regelverk de forvalter, er bemyndiget til å gi uttalelser eller gjøre vurderinger i en byggesak. Det vil f.eks. være Riks- eller Byantikvaren for endringer på gamle bygg, Arbeidstilsynet for arbeidslokaler etc.

4 Oversikt over funnene

Totalt har kartleggingen identifisert 93 kilder som kan være relevant for byggesak. Det gjenspeiles i oversikten at definisjonen av *datasett* ikke er helt åpenbar. På en enkel måte kan man si at et datasett er *en samling med data som det gir mening å snakke om som en enhet*. Dette er ingen presis beskrivelse, men vi har likevel valgt å gå for denne, med tanke på breddefokuset og for å forsøke å favne om alle potensielt relevante kilder og faktagrunnlag.

4.1 Formater

Formatet som de forskjellige datasettene tilbys i, kan deles kvalitativt inn i tre:

1. Strukturerte og standardiserte formater (SOSI, GeoJSON, GML, KML, Esri Shape og andre geoformater)
2. Programmeringsgrensesnitt (Web services, WMS/WFS, REST og SOAP API)
3. Ustrukturerte, «menneskelesbare» formater (PDF, Excel, bildeformater, html, kart kun til innsyn osv.)

Prosentvis fordeler datasettene seg på denne måten: 57 % av kildene tilgjengeliggjøres i ett eller annet strukturert og maskinlesbart format (1), 63 % av datasettene tilbys gjennom et programmeringsgrensesnitt (2), og 25 % av de registrerte kildene finnes kun i «menneskelesbare» formater (3). Det er viktig å være oppmerksom på at det i mange tilfeller er overlappende mengder. Mange dataforvaltere tilbyr de samme datasettene både gjennom innsynsløsninger, web-tjenester og som rådata til nedlasting.

De standardiserte datautvekslingsformatene for kartdata som brukes i størst omfang er SOSI (norsk standard for geografisk informasjon), GeoJSON, GML, KML (Google) og Esri Shapefile. Det finnes konverteringsverktøy mellom de forskjellige formatene. Imidlertid er de fleste geoformatene (og særlig Shape) mindre komplekse enn SOSI, og man kan dermed risikere å miste informasjon i konvertering fra SOSI (f.eks. blir topologier ofte forflatet til enkelte objekter, kurver interpolert til en rekke av rette linjer, og man kan miste lange attributtnavn og norske og samiske tegn).

4.2 Tilgjengelighet

Av alle kartlagte datakilder er 72 % åpent tilgjengelige. Med åpent menes her enten at datasettene ganske enkelt kan lastes ned med et tastetrykk, eller at det kreves at brukeren registrerer seg, men tilgang gis (mer eller mindre) automatisk.

18 % av kildene er «tilgjengelige med begrensninger». Med dette menes at datasettene enten delvis er skjermet for allment innsyn (f.eks. på grunn av sikkerhetsvurderinger eller fredningshensyn), kun tilgjengelig for den offentlige forvaltningen (såkalte «gule data» i henhold til Difis trafikkløssystem), bare deler av datasettet er publisert, eller at de er tilgjengelige, men upubliserte.

Kartleggingen har også forsøkt å avdekke hvilke datasett som ikke er tilgjengelige for noen eksterne parter i dag, men som har potensiale for å bli det med tiden. Slike funn utgjør 9 % av totalen. Grunnene som oppgis for at disse ikke er tilgjengeliggjort er enten at man ikke har visst om behovet, dataene ikke holder høy nok kvalitet, eller at de av ikke ennå er ferdigstilt.

4.3 Bruk av lisenser

56 % av datakildene er tilgjengeliggjort under en åpen lisens. Det er to slike lisenser som er i bruk i funnene: Norsk lisens for offentlige data (NLOD) i klart størst grad og Creative Commons i mindre grad. NLOD er en standardisert offentlig norsk lisens. Hensikten med lisensen er å gjøre det enklere for dataeiere å presisere vilkårene for bruk av datasettene de publiserer og å bruke data fra forskjellige kilder sammen. Kort oppsummert sier lisensen at du har lov til å bruke informasjonen til det du vil, men at du må oppgi hvor du har hentet informasjonen fra. Lisensen gjør det klart at sensitive og taushetsbelagte opplysninger ikke kan tilgjengeliggjøres under lisensen, og at dataeier ikke har ansvar

verken for informasjonens kvalitet eller for hva informasjonen brukes til. Creative Commons Navngivelse er en internasjonal lisens som gir lov til å dele (kopiere, distribuere og spre data i et hvilket som helst format) og bearbeide (endre og bygge videre på) materialet til et hvilket som helst formål (inkludert kommersielle) på vilkår om korrekt kreditering, henvisning til lisensen og indikasjon om at endringer er gjort.

De øvrige kildene har ikke spesifisert lisens.

5 Variasjon i tilnæringsmåte

Det er stor variasjon i hvor langt de ulike myndighetene har kommet når det gjelder å tilgjengeliggjøre data. Blant de virksomhetene som har jobbet mest med dette er Statens vegvesen, Miljødirektoratet, Kystverket og Norges geologiske undersøkelse (NGU). Her har man valgt å rette spesielt stort fokus på tilgjengeliggjøring av offentlige data, med tydelig forankring i ledelse og strategi. Disse aktørene har satt av egne ressurser til å sørge for at datasett publiseres i en rekke standardiserte formater og tjenester, og de er aktive deltakere i store nasjonale samarbeidsprosjekter som Geovekst, Norge Digitalt og Geonorge.

Eksempelvis ble Statens vegvesen i 2013 kåret til «Norges mest åpne etat» på Difis digitaliseringskonferanse. Vegvesenet ser stor nytte i å gjøre digitale datasett tilgjengelige for å få andre i markedet til å utvikle gode elektroniske tjenester, samtidig som de er opptatt av at brukerne kan melde inn feil på data direkte knyttet opp mot Vegvesenets databaser. De fremhever spesielt at feilmeldingstjenesten bidrar til god dialog med brukerne og sørger for stadig forbedring av dataene.

Andre fagmyndigheter har ikke kommet like langt med å tilgjengeliggjøre sine data. Det kan virke som om tilgjengeliggjøring av data ikke er et tema i noen av virksomhetenes strategiske satsning. Trass dette opplever man at det på fagnivå finnes et stort engasjement, der man ser potensialet i å publisere egne datasett til viderebruk. Eksempel på dette er Jernbaneverket, som har en rekke datasett av potensiell nytte for byggesaksprosessen, men disse er ikke åpne, eller kun synlige i «menneskelesbare» formater. Jernbaneverket har utviklet en omfattende kartløsning med en rekke datasett som kun ment til intern bruk. Den er ikke publisert, men likevel tilgjengelig, så lenge man kjenner til adressen (løsningen er omtalt som «indirekte åpen» i datakildeoversikten i vedlegg 1). Jernbaneverket bidrar på sin side til mange andre databaser der data fra flere leverandører sammenstilles, for eksempel Nasjonal database for grunnundersøkelser (NADAG) og skredatlas.nve.no.

Igen andre opplyser at de ikke har noen registre med data i maskinlesbare formater. Dette er typisk virksomheter som ikke har behov for faktagrunnlag i form av databaser for å ta beslutninger, men bygger på regelverket, veiledninger til disse, erfaringsgrunnlag og skjønn. Arbeidstilsynet er et slikt eksempel. På spørsmål om hva man kunne tenke seg i fremtiden ville være av allmenn nytte sett fra et selvbetjeningsperspektiv, pekes det på tilgjengeliggjøring av erfaringsgrunnlag fra tidligere saker som kan sammenkobles på en meningsfull måte med den enkelte problemstilling.

Inntrykket er at myndighetene generelt viser stor interesse for deling av informasjon. Mange gode initiativer er igangsatt for å gjøre flere offentlige data tilgjengelig til gjenbruk. Spørsmålene som ble stilt under kartleggingen vakte oppmerksomhet og entusiasme, noe som tyder på at myndighetene ser potensialet i lettere tilgjengelige data og muligheten til å redusere mye tids- og ressursbruk som i dag går med til å lete og gjenskape.

5.1 Problemer og hindringer

Det kan være flere grunner til at ikke alle datasett er åpent tilgjengelige i dag. En av dem er opphavsrettslige og dermed økonomiske hindringer. Eksempel på dette er resultatene av boringer og grunnundersøkelser. Dette er data som produseres på bestilling av Jernbaneverket, Statens vegvesen, NGU, kommuner, bedrifter og privatpersoner. Boringene og undersøkelsene utføres av private konsultantselskaper som hevder opphavsretten til resultatene og ønsker å gjenbruke dem gjennom videresalg. Dette strider mot prinsippet om at arbeid som utføres på bestilling av det offentlige også skal være åpent tilgjengelig for det offentlige. NADAG gir en samlet oversikt over hvilke undersøkelser som er utført, men det er forskjell på graden av tilgjengeliggjøring fra de forskjellige dataleverandørene. Statens vegvesen har besluttet å publisere data fra grunnundersøkelser uavhengig av fremtidige eierskaps- og rettighetskonflikter, mens andre dataleverandører er foreløpig mer tilbakeholdne med publisering.

Manglende strategisk forankring i virksomhetsledelsen kan være en årsak til at datasett ikke publiseres, som nevnt innledningsvis i kapitlet. En annen sak er at det ikke nødvendigvis eksisterer

motvilje eller hindringer mot deling av data, man har rett og slett ikke tenkt på at det kan være til nytte for noen. Et eksempel er Sivilforsvarets register over tilfluktsrom. Intervjuet med Sivilforsvaret avdekket at det ikke er noe i veien for å dele dataene over eksisterende tilfluktsrom, dersom det kunne være til nytte for noen.

En teknologisk hindring er at SOSI-formatet oppleves som noe tungvint. Det er svært omfattende både å produsere og bruke, og flere kommuner etterlyser data i Esri Shape eller andre geoformater. Disse er imidlertid mye mindre komplekse en SOSI, og man risikerer å miste vesentlig informasjon (som lange attributtnavn og norske og samiske tegn) ved konvertering fra for eksempel SOSI til Shape.

Data som er skjermet for offentligheten av diverse årsaker – for eksempel sikkerhetshensyn og sabotasjefare rundt storulykkanlegg, personvernensyn ved visning av bilder av interiør i kulturminner – blir forhindret fra å tilgjengeliggjøres.

Et område det ser ut til å være en del usikkerhet rundt i etatene er hvordan å kombinere åpenhet og sikkerhet. Det er en del informasjon som ikke bør offentliggjøres uten videre på grunn av rikets sikkerhet, personvernensyn eller faren for tap av utrydningstruede arter. Det er vårt inntrykk at å publisere disse dataene i anonymisert form eller på andre måter gi brukerne indikasjoner på etatens interesser, i liten grad er diskutert eller avklart i mange av sektormyndighetene.

En annen begrensning er at noen av sektormyndighetene er pålagt et inntjeningskrav og krever betaling for sine data. Dette gjelder bl.a. Matrikkelen og en del andre kartdata fra Statens kartverk. I den siste tiden har Kartverket frigitt mange datasett og det betyr at denne utfordringen ikke lenger er så fremtredende som tidligere.

Til slutt, den hyppigste grunnen som oppgis til manglende publisering er at dataene ikke holder høy nok kvalitet. I flere tilfeller ble det opplyst at man ikke ønsker å tilgjengeliggjøre datasettet før det eksempelvis kunne sies å være landsdekkende eller før kvaliteten på dataene kunne heves. Under avsnitt 3.3 nevnte vi at eksponering av data også har den fordel at brukernes tilbakemeldinger kan brukes til å korrigere datakvaliteten. Flere dataforvaltere tilbyr tjenester der de oppfordrer publikum til å melde fra om sine funn, mangelfulle eller feilaktige data, såkalt *crowdsourcing*. Eksempelvis kan hvem som helst legge inn egendefinerte kulturminner gjennom Riksantikvarens nettjeneste Kulturminnesøk, skader i veien gjennom Statens vegvesens tjeneste «Fiksvegdata» og varsle om nye funn gjennom Artsdatabankens meldingsløsning. Denne måten å berike data på har sine fordeler i at man høster tilbakemeldinger direkte fra brukerne og kan til en viss grad danne seg et bilde av den faktiske bruken, populariteten og forbedringspotensialet på den tilgjengeliggjorte informasjonen. Samtidig er ikke alle erfaringene med *crowdsourcing* utelukkende positive. Det å skille verdifulle bidrag fra de ubrukelige er en ressurskrevende oppgave som kan stjele fokus fra å jobbe med dataforbedring gjennom autoriserte redaktører. Dessuten kan publisering av publikums bidrag uten noen form for redaksjonell, faglig kontroll skape mye støy i presentasjonen av datasettet.

Av de kartlagte datasettene er det 19 av totalt 93 datasett som kan påvirkes av publikum i en eller annen form. I tellingen har vi tatt med alle former for oppfordring til brukergenerert informasjon, også de tjenestene som består av enkle skjemaer til å sende inn meldinger, slik som «Varsle Kystverket» og tilbakemeldingsskjemaene på NGUs innsynskart.

5.2 Kommunikasjon med kommunene

Flere myndigheter uttrykker at det eksisterer et potensial for forbedret kommunikasjon og økt samhandling mellom kommunene og sektormyndighetene. Det er i dag lite automatikk i innhenting av faktagrunnlag og uklare rammer rundt samarbeidsprosessene. Mye av arbeidet er papirbasert og prosessen er lang og møysommelig. Gjensidig tillit til riktig bruk av dataene og gode elektroniske tjenester vil kunne gi store besparelser, både på tid og ressurser.

Et eksempel er utfordringer knyttet til håndteringen av SEFRAK-bygninger. Riksantikvaren har forvaltningsansvaret for det såkalte SEFRAK-registeret over bygninger (i hovedsak eldre enn år 1900) som regnes som kulturminner, men som ikke er fredet etter Kulturminneloven. Kildedataene til SEFRAK ligger fysisk i Matrikkelen som forvaltes av Statens kartverk, slik at eierskapet for datasettet i

realiteten er delt. Dette resulterer i at oppdatering av registeret er tungvint og man har opparbeidet seg et visst etterslep på datakorrigeringer. I tillegg er kanskje det mest verdifulle materialet, bildedokumentasjonen, ikke tilgjengelig i Matrikkelen, men oppbevares i mange tilfeller fysisk i kommunenes arkiver. Kommunene er flittige brukere av disse dataene og har selv et stort behov for å oppdatere og hente ut registreringer. Derfor vedlikeholder mange kommuner sine egne registre (i forskjellige formater) av SEFRAK-bygninger som ikke nødvendigvis er ført i det sentrale SEFRAK-registeret. På sikt vil denne situasjonen skape forvirring og man vil miste oversikten over SEFRAK-bygninger. Det er for tiden god dialog mellom Kartverket og Riksantikvaren for å forsøke å finne en løsning før problemet eskalerer.

Et annet eksempel er administrasjonsgrensene for reindriftsområder som avviker fra de kommunale administrasjonsgrensene. Man opplever at saksbehandling blir i visse tilfeller unødvendig komplisert og tidkrevende fordi områdegrensene og bruken av dem ikke er godt nok synliggjort og samholdt med andre data. Ved å løfte frem og integrere data fra reindriftsforvaltningen i byggesaksbehandlerens fremtidige elektroniske løsninger, ville man unngått misforståelser knyttet til grenser og ansvarsområder, og saksbehandlingstiden ville blitt redusert betraktelig.

Støysonekart fra Jernbaneverket er nok et eksempel på et datasett som muligens kan bidra til smidigere samhandling mellom kommuner og myndigheter. I dag er det store avvik på hvordan støysonekart blir tolket av byggesaksbehandlere og hvor plagsom mennesker oppfatter den aktuelle støyen. Tyding av kartene krever høy kompetanse som man ikke kan forvente at alle saksbehandlerne besitter. Ved bygging av nye boliger langs jernbanelinjen er det en tilbakevendende problemstilling er at boligkjøpernes forventninger ikke står i stil med det produktet de får levert med tanke på støybelastning. Tillatelse til bygging gis relativt nær jernbanelinjen og i visse tilfeller uten grundig nok vurdering av støybelastningen både innendørs og utendørs. På sikt vil dette påvirke boligens verdi. Økt forståelse for og bedre synliggjøring av akustiske data til byggesaksbehandlere ville senket boligkjøpernes forventninger til (mangel på) støy, men til gjengjeld gitt et mer sannferdig bilde av situasjonen.

Et viktig ledd i myndighetenes kommunikasjonen med kommunene innenfor området byggesak er etableringen av Det offentlige kartgrunnlaget (DOK). DOK vil være et samlingssted for offentlige og kvalitetssikrede geografiske data som er tilrettelagt spesielt for plan- og byggesaksarbeid. En rekke av datasettene vi har kartlagt er allerede med i DOK eller kandidater til publisering i DOK ved et senere tidspunkt. På det generelle planet finnes større statlige initiativer for tilgjengeliggjøring av data. Det mest sentrale er Data.norge.no, som er et register og datahotell for åpne offentlige data i Norge. Her publiseres data i mange forskjellige maskinlesbare formater og temaer. Det stilles lave krav til standardisering og spesifikasjoner. I tillegg til dette har Statens kartverk fått i oppdrag å lage en nasjonal geoportal: Geonorge. Målet er at denne skal være et samlingspunkt for publisering, oppslag og tilgang til all norsk geografisk informasjon innen 2018.

Denne kartleggingen omfatter ikke en detaljert gjennomgang av fordelingen av datasett mellom data.norge.no og GeoNorge.

6 Konklusjon

Dialogen med de respektive myndighetene har avdekket at det oppfattes å være et potensial for forbedring av effektivitet og særlig kommunikasjon med kommunene gjennom økt deling av informasjon. Det er gjennomgående positiv innstilling til økt samhandling gjennom tilgjengeliggjøring av data slik DiBK legger opp til gjennom ByggNett-strategien.

Kartleggingen har vist at det finnes veldig mange datasett som er eller kan være til nytte i byggesaksprosessen hos de forskjellige myndighetene. Oversikten viser et landskap med stor kompleksitet, mange formater og ulik kvalitet på datasettene.

Det er gledelig at det pågår arbeider for samordning og tilrettelegging av åpne data på flere hold i dag. Data.norge.no, Geonorge og Det offentlige kartgrunnlaget er gode initiativer med store målsettinger, men det gjenstår fremdeles en del arbeid i avklaringer av formater, lisenser, spesifikasjoner og eierforhold. Det legges her mye vekt på geografisk informasjon, men det finnes også en rekke datasett som ikke er stedfestet. Disse ser ut til å bli noe undervurdert i de nasjonale satsingene. Ikke minst ligger det et stort potensial i å publisere adresserbare juridiske definisjoner for å gjøre beslutningene i den enkelte sak mer transparente og sporbare.

Behovet for åpne data til bruk i selvbetjeningsløsninger – slik visjonen om Sjur beskriver det – er ikke dekket i dag. For at visjonen skal kunne bli virkelighet, må det stilles krav til at datasett publiseres i standardiserte formater, er pålitelige, oppdaterte og relevante. Det vil være en enorm besparelse i byggesaksprosessen dersom kontekstbasert grunnlagsdata kunne blitt løftet frem i de initielle leddene, fra tiltakshaver starter sin vurdering av å sende inn en søknad til en saksbehandler sender saken til høring. En naturlig anbefaling er å etablere et felles *nav* for de relevante datasettene til bruk i byggesaksprosessen, der informasjonen fra ulike etater og faginstitusjoner lenkes gjennom semantiske modeller. På denne måten er ikke data bare åpne, men også lenket sammen gjennom stabile og entydige identifikatorer. Det er mye inspirasjon å hente i prosjektet Semicolon som nettopp tar for seg organisatorisk og semantisk samhandling i og med offentlig sektor. Denne tilnærmingen sikrer at endringer i data, regelverk, struktur og prosesser alltid vil oppdateres automatisk i sin kontekst takket være de entydige identifikatorene og lenkingen.

Et eksempel på en slik satsing er Utdanningsdirektoratets «Grep»-tjeneste der man tilbyr læreplaner, kodeverk og tilbudsstruktur beskrevet i en delt semantisk modell i stedet for hundrevis av *menneskelesbare* dokumenter. Man kan i fremtiden tenke seg en lignende løsning på utfordringen med det mangfoldige regelverket og det enorme informasjonsbehovet rundt byggesaksprosessen. Til dette trengs fellestjenester som indekserer og sammenstiller de sentralt tilgjengeliggjorte definisjonene og modellene, sammen med data om autoritet og faktisk bruk.

Vedlegg

Vedlegg 1: Totaloversikt i tabellform

Vedlegg 2: Nettressurser fra Miljødirektoratet

Vedlegg 3: Intervjuguide