

Utredning om egnede standarder for definisjonsarbeid i offentlig sektor

Sluttrapport
25. januar 2010

David Norheim, Computas,
Lasse Udjus, Karde
Henrik Smith-Meyer, Det Norske Veritas

1	Forord.....	3
2	Definisjonsarbeid.....	4
2.1	Standardiseringsrådets generelle kriterier til forvaltningsstandarder	4
2.1.1	Hva vi legger i en begrepsdefinisjon?.....	4
2.1.2	Definisjonsprosessen.....	5
2.2	Kriterier for støtte til definisjonsarbeid.....	7
3	Behovsanalyse.....	8
3.1	Bakgrunn.....	8
3.2	Deltakere i prosessen	9
3.3	Interessenter og målgruppe for behovsanalysen	10
3.4	Om prosessen	10
3.4.1	Arbeidsmøte 1 - 8. oktober 2009	10
3.4.2	Arbeidsmøte 2 - 19. november 2009.....	11
3.5	Definisjoners anvendelsesområder og avgrensninger.....	11
3.5.1	Regelverk og rammebetingelser	13
3.5.2	Intern forvaltning	17
3.5.3	Utteksling av informasjon.....	18
3.5.4	Innsyn og tilgjengeliggjøring.....	20
4	Kriterier	22
4.1	Krav til for definisjonsarbeid i offentlig sektor	22
4.1.1	Allmenn forståelse og språk.....	22
4.1.2	Kontekst	22
4.1.3	Definisjonens innhold, nøyaktighet og kvalitet	23
4.1.4	Strukturering av begreper	24
4.1.5	Refererbarhet.....	24
4.1.6	Endringshåndtering og sporbarhet	25
4.1.7	Gjenbrukbarhet og harmonisering	25
4.2	Presiseringer og avgrensninger	25
5	Vurdering av standarder.....	26
5.1	Avgrensninger.....	26
5.2	Standardene og anvendelser.....	30
5.3	Om standardiseringsorganisasjoner og initiativene	31
5.4	Kort om de vurderte standardene og initiativene.....	32
5.5	Evaluerings mot kriteriene	34
6	Anbefalinger for videre arbeid.....	36
6.1	Anbefalinger relatert til anvendelsesområdene for definisjoner.....	36

1 Forord

I vårt arbeid med å finne, kartlegge og evaluere standarder relevante for denne rapporten har vi søkt råd og hjelp hos en rekke personer i en rekke offentlige organisasjoner. Alle har bidratt med iver og innsikt. Vi ønsker derfor å takke følgende personer og organisasjoner for bidrag til utredningen: Anne Gro Hustoft, Jenny Linnerud, Line Richardsen, Ragna Bruland, Jon Arve Risan, John Martin Furset, Kristian Bergem, Endre Grøtnes, Espen Seland, Jostein Ven, Gyrid Espeland, Janneke Solem, Ellen Hestenes, Audun Aanæs, Are Mjøltnes, Arild Jansen, Øystein Aanrud, Ståle Norum Engen, Morten Hagedal, Tormod Kristiansen, Dag Wiese Schartum, Pia Jøsendal, Gunnar Coll, Per Myrseth, Frode Ettesvoll, Geir Jevne, Terje Grimstad, Espen Slotvik, Magne Valen-Senstad, Johan Klüver, Håvard Hjulstad, Steinar Skagemo og Jon Leirdal.

Rapporten er strukturert som følger

- Vi starter med å beskrive forutsetningene for en felles forståelse av definisjonsprosessen. Vi legger her også frem en liste med begreper sentrale i definisjonsarbeid.
- Definisjonsprosessen blir så satt i sammenheng med i hvilke situasjoner manglende begrepsdefinisjoner kan føre til problemer.
- Deretter fremlegger vi en rekke kriterier standarder for definisjonsarbeid bør vurderes opp mot, samt at vi trekker noen avgrensninger i forhold til hvilke standarder og initiativer som vil bli vurdert.
- Kriteriene kategoriseres og anbefales standarder på grunnlag av evalueringene.
- Innstillingene til de anbefalte standardene baserer seg på erfaringer og vurderinger av hva som er realistisk å forvente, sammenfallende med forventninger til åpenhet, tilpassning og anvendelse såvel hos brukere som informasjons- og programvareleverandører.

Vi vil tilslutt takke Difi for anledningen til å bidra på dette feltet som vi anser som svært spennende og fremtidsrettet.

2 Definisjonsarbeid

Denne rapporten er leveransen til DIFIs utlysning av “Utredning av egnede standarder for definisjonsarbeid i offentlig sektor”. Utredningsarbeidet har fulgt Standardiseringsrådets utredningsprosess vist i figur 1.



Figur 1 - Standardiseringsrådets utredningsprosess

I dette kapittelet ønsker vi å sette fokus på hva vi oppfatter som problemstillinger innen definisjonsarbeid, og kommer så tilbake til behovsfasen i Kapittel 3.

2.1 *Standardiseringsrådets generelle kriterier til forvaltningsstandarder*

Standardiseringsrådets prosessdokument har en rekke generelle kriterier som skal legges til grunn i vurderingen om en standard er egnet som forvaltningsstandard i offentlig sektor. Kriteriene kan sorteres i fire kategorier:

- **Formålet med standarden.** Kriterier rettet mot hensikten til standarden og bruksområdet standarden er ment for.
- **Proessen rundt utvikling og forvaltning av standarden.** Kriterier som adresserer beslutningsprosessen, anerkjennelse og åpenhet, samt tilgjengelighet og dokumentasjon og IP rettigheter.
- **Standardens aksept og utbredelse i markedet.** Kriteriene adresserer modenhet, implementasjon, adopsjon o.l.
- **Tilfredsstiller standarden offentlige virksomheter og deres brukeres behov?** Krav knyttet til brukssituasjonen, kvalitetskrav o.l. Disse kriteriene blir ytterligere detaljer gjennom de neste kapitlene.

Det komplette settet av kriterier er gjengitt i bilag A – ”kriterier for definisjonsarbeid”.

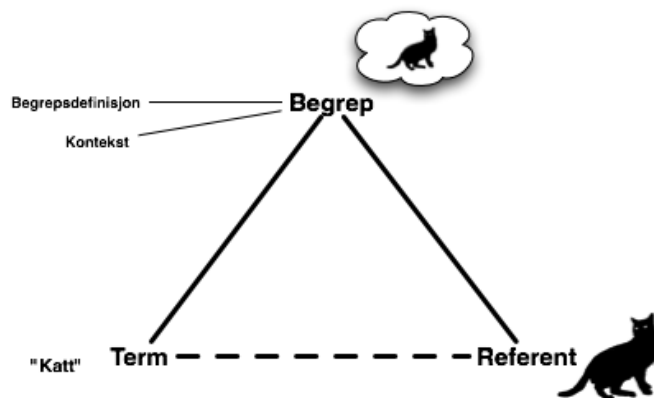
2.1.1 Hva vi legger i en begrepsdefinisjon?

Nedenfor er en forklaring av noen av de sentrale begrepene i forbindelse med definisjonsarbeid.

Tabell 1 - Begrepsforklaringer

Begrep/Term	Forklaring
Referent	Det en term og begrep omtaler
Term	Ord eller uttrykk brukt for referenten
Begrep	Klart avgrenset og allmenngyldig forestilling brukt om referenten
(Begreps-)definisjon	Det å definere; nøyaktig begrepsforklaring. En definisjon er en kort formell tekst som beskriver meningen av en term
Kontekst	En sammenheng begrepene må forståes. F.eks. et fagområde. Begrepet "Buss" kan f.eks. forståes svært forskjellig innen IKT og Transportområdet.
Begrepsapparat	En samling av begreper i en kontekst

Ulike fagmiljøer, terminologi, lingvistik, filosofi, kunnskaps-representasjon, biblioteksfag og logikk m.fl., har noe ulikt utgangspunkt rundt begrepene over¹. Figur 2 illustrerer én sammenhengen mellom term, begrep og referent kalt det semiotiske triangel. Vi velger å tar utgangspunkt i denne, som vist under, som vi oppfatter som den mest samlende for disse fagmiljøene.



Figur 2 - Charels Kay Ogdens "Semiotic Triangle" med tilknyttet begrepsdefinisjon og kontekst.

2.1.2 Definisjonsprosessen

Vi oppfatter definisjonsarbeid i offentlig sektor som en prosess, og ikke et prosjekt², ettersom begreper vil være i konstant endring. Når standarder for definisjonsarbeid skal vurderes er det nødvendig å påpeke at standardene dekker ulike deler av en slik prosess. I

¹ For eksempel opererer lingvister med Synsets som et sett av synonymer som har lik mening. Dette tilsvarer Begrep i figuren.

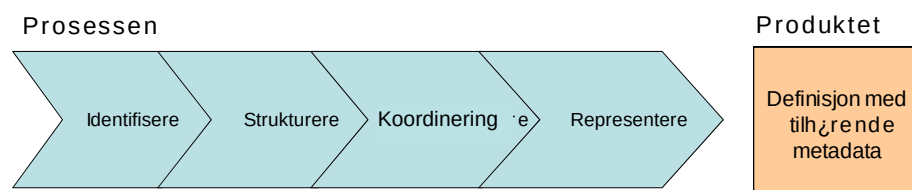
² Struktureringen av et begrepsapparat blir innen enkelte fagkretser (ISO 704) ofte kalt terminologiarbeid, og betegnes her ofte som et prosjekt.

figur 3 og Figur 4 illustrerer vi en prosess som en ramme for å sammenlikne standardene, men den er ikke ment å være en beskrivende prosess for alt definisjonsarbeid³.

Før man har standarder til støtte er det behov for å kartlegge omfang og krav til definisjoner for en bestemt anvendelse av definisjonen. Anvendelsene blir kartlagt i kapittel 3.5 med de spesielle kriteriene dette stiller til begrepsdefinisjoner.

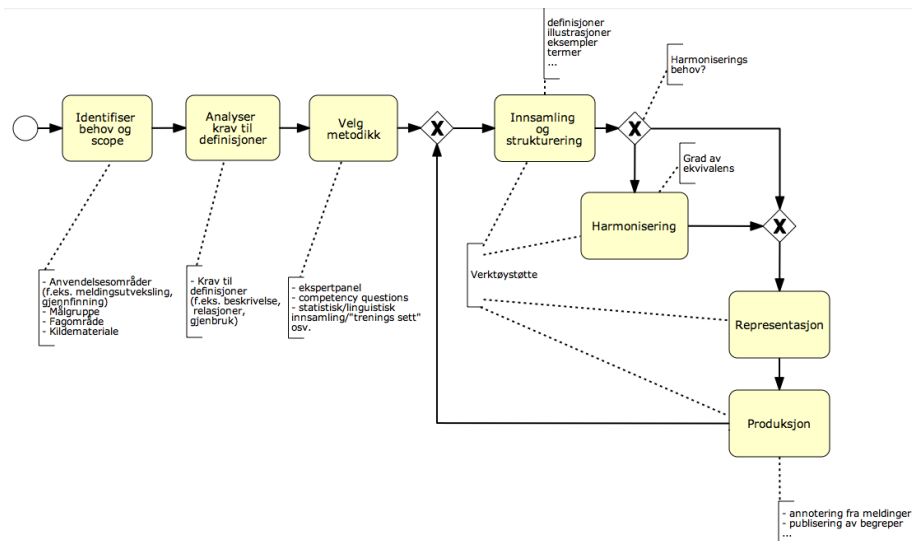
Vi fokuserer på å vurdere relevante standarder innenfor de fire del-prosessene for hvert bruksområde:

- Metodikk for innsamling av begreper
- Innsamling, strukturering av begreper og begrepsdefinisjoner
- Koordinering (gjenbruk og harmonisering av begreper)
- Representasjon av begreper og begrepsapparatet



Figur 3 - Arbeid med definisjoner: prosess og produkt. Produkt er metadata som knyttes til et begrep og begrepsdefinisjonen.

³ Det er ikke i prosjektets mandat å detaljere denne prosessen. Difi utarbeider en slik prosess parallellt med dette prosjektet.



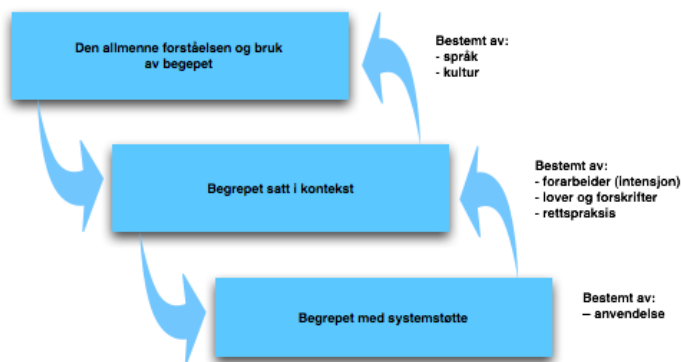
Figur 4 – Delprosessene for definisjonsarbeid er satt i en sammenheng med relevante steg der det finnes standard-støtte. Figuren må kun oppfattes som en skisse på delprosesser som inngår i definisjonsarbeid, og er ikke ment som en "fasit" på definisjonsprosessen.

Gjennom eksempler vil vi forsøke å illustrere hvordan en forbedret definisjonsprosess vil kunne gi gevinster, og hva som i det enkelte tilfelle er spesielle hensyn som må ivaretas for å oppnå ønsket effekt.

2.2 Kriterier for støtte til definisjonsarbeid

For å forstå de ulike behovene for definisjoner har arbeidsgruppen forsøkt å komme opp med eksempler der en god anvendelse av definisjoner vil gi klare gevinster. Ut fra eksemplene ønsker vi å avdekke hva som karakteriserer en god definisjon i det aktuelle anvendelsen.

Figur 5 illustrerer tre gjensidig påvirkende nivåer av en godt fundert definisjon. Først og fremst bør definisjonen være forankret i språk og kultur. Begrepet og definisjonen bør videre forstås i den kontekst etaten representerer. Sist, men ikke minst, må begrepets anvendelse forstås slik at det representeres riktig i systemer.



Figur 5 – Generelle krav til Begrepsdefinisjoner og gjensidig påvirkning

3 Behovsanalyse

Oppdragsgiver har tidlig gjort det klart at det overordnede målet for denne utredning om standardisering er samhandling og rettssikkerhet. Vi har derfor tatt tak i problemet fra behovssiden. Hovedspørsmålet vårt blir da:

I hvilke situasjoner er mangelfulle definisjoner blant årsakene til dårlig eller fordyrende samhandling eller mangelfull rettssikkerhet?

Vi starter denne behovsanalysen med en kartlegging av behovet for gode definisjoner i offentlig sektor.

3.1 Bakgrunn

Utredningen av standarder for definisjonsarbeid var i utgangpunktet ikke avgrenset til å gjelde spesifikt en bestemt type definisjoner eller tilhøre bestemte fag- eller forvaltningsområder. Med utgangspunkt i tidligere utredninger, som har hatt et tyngdepunkt på IT-siden i å vurdere behovene for eksplisitte definisjoner, datadefinisjoner og metadata, ble det imidlertid presisert at denne utredningen spesielt også skulle vurdere de juridiske og forvaltningsspesifikke behovene i arbeidet med definisjoner.

For å ikke favne for vidt valgte arbeidsgruppen i samråd med oppdragsgiver å fokusere på et utvalg etater, som alle har utfordringer knyttet til samhandling, som går utover rapporteringsplikt.

3.2 Deltakere i prosessen

Tabell 2 - Deltakere i prosessen

Organisasjon	Deltaker	Workshop 1 (8.oktober 2009)	Workshop 2 (19.november 2009)
Aktører for vurdering av behovssiden:			
SSB	Anne Gro Hustoft	x	x
SSB	Jenny Linnerud	x	
KS	Line Richardsen	x	
Statens pensjonskasse	Ragna Bruland	x	
Statens pensjonskasse	Jon Arve Risan		x
UDI	John Martin Furset	x	
Difi	Kristian Bergem	x	x
Difi	Endre Grøtnes	x	
KITH	Espen Seland	x	
KITH	Jostein Ven	x	
Mattilsynet	Gyrid Espeland	x	x
Mattilsynet	Janneke Solem	x	
Mattilsynet	Ellen Hestenes	x	x
Justisdepartementet	Audun Aanæs	x	
Utdanningsdirektoratet	Are Mjøltnes	x	
UiO/AFIN (RneF)	Arild Jansen	x	
Bærum kommune	Øystein Aanrud	x	x
Brønnøysundregistrene	Ståle Norum Engen	x	
Skattedirektoratet	Morten Hagedal	Særmøte 28.oktober	x
NAV	Tormod Kristiansen		Særmøte 27.november
UiO/AFIN	Dag Wiese Schartum	x	Særmøte 17.november
Computas	David Norheim	x	x
Computas	Pia Jøsendal	x	x
Computas	Gunnar Coll	x	
DNV / Semicolon	Henrik Smith- Meyer	x	x
DNV / Semicolon	Per Myrseth	Særmøte 28.oktober	x
Karde / Semicolon	Lasse Udjus	x	x
More	Frode Ettesvoll	x	x
DNV / Seres	Geir Jevne	Særmøte 28.oktober	

Organisasjon	Deltaker	Workshop 1 (8.oktober 2009)	Workshop 2 (19.november 2009)
Karde / Semicolon	Terje Grimstad	Særmøte 28.oktober	
Brønnøysundregistrene / Semicolon	Espen Slotvik	Særmøte 28.oktober	
Difi	Steinar Skagemo		Særmøte 8.desember
Aktører på standardiseringssiden:			
DNV	Magne Valen- Senstad		Særmøte 28.november
DNV	Johan Klüver		Særmøte 28.november
Standard Norge	Håvard Hjulstad		Særmøte 8.desember

3.3 Interessenter og målgruppe for behovsanalysen

Selv om behovsanalysen er gjort på bakgrunn av de ovenfor listede deltakernes deltakelse vurderes likevel behovene å være representative for også andre deler av offentlig forvaltning.

Det vil være opp til oppdragsgiver, Difi ved Sekretariatet for Standardiseringsrådet, å vurdere hvor bredt gyldighetsområde et eventuelt valg av standarder for definisjonsarbeid måtte gjelde.

3.4 Om prosessen

Behovsanalysen har blitt gjennomført ved at utvalgte offentlige etater ble invitert til å delta i to workshops for å gjennomgå deres behov for støtte i arbeidet med definisjoner.

3.4.1 Arbeidsmøte 1 - 8. oktober 2009

Første arbeidsmøte handlet om å presentere noen grunnleggende problemstillinger rundt definisjoner for deltakende etater, og videre invitere alle til å diskutere hvilke utfordringer den enkelte etat stod overfor.

Det ble tidlig presisert at forsamlingen var bredere representert på IT-siden enn på juridisk og forvaltningsfaglig side. Også de juridiske og forvaltningsfaglige sidene bør være godt representert i diskusjoner om definisjonsarbeid. Denne feilrepresentasjonen oppfattes som et generelt problem i denne typen behovsanalyser. Dette reflekterer også hvem i forvaltningen denne problemstillingen har hvilt på over tid.

Noen hovedkonklusjoner fra arbeidsmøte er:

1. Det er behov for et begrepsapparat for å bli tydelige på begrepene vi benytter (f.eks. definisjon, begrep, term etc.).

2. Det er viktig å ha et klart forhold til hva som defineres. Spesielt bør det vurderes hvorvidt man skal ha veldig presise definisjoner innenfor dynamiske fagområder.
3. Definisjoner er ofte gitt ved flere forhold. For eksempel definisjonen av ”samboer”, der forståelsen av begrepet må tolkes i kontekst av i hvilken sammenheng (les: hvilken lov/forskrift) begrepet brukes. Det må være mulig å referere på tvers av definisjoner, slik at en definisjon kan bestemmes ved koblingen mellom ulike definisjoner.
4. Det mangler IT-støtte til arbeid med lovgivning, og slike verktøy kan åpenbart ha viktige funksjoner i arbeidet med begrepsanalyse og -harmonisering (ref Schartum).
5. Det er behov for bedre støtte til forvaltning av definisjoner på tvers av fagområder (anvendelse av definisjoner i lover mv.). Viktig å formidle definisjoner mellom ulike deler av forvaltningen (uansett definisjonsmessig likhet/ulikhet).
6. Det finnes noen gode eksempler på godt og tett samarbeid mellom juridisk/forvaltningsfaglig side og IT-siden. Spesielt har man god erfaring med dette hos Mattilsynet.

3.4.2 Arbeidsmøte 2 - 19. november 2009

Til det andre arbeidsmøtet hadde arbeidsgruppen gjennomført enkeltmøter med Skattedirektoratet (som var forhindret fra å delta på arbeidsmøte 8. oktober) og med SERES- og Semicolonprosjektet. Til dette arbeidsmøte hadde arbeidsgruppen systematisert ulike aspekter ved bruk av definisjoner, med en liste over ulike behov og kriterier som må legges til grunn i vurdering av hvilke standarder som kan understøtte disse. Det ble også gitt en presentasjon av hvordan arbeidsgruppen vurderer ulike standarder opp mot de aktuelle kriteriene. Arbeidsgruppen har også avholdt et særmøte med NAV i etterkant av den andre workshopen.

De neste avsnittene gir en gjennomgang av ulike anvendelser og de spesielle krav det stiller til definisjonsarbeidet.

3.5 Definisjoners anvendelsesområder og avgrensninger

Arbeidsgruppen har ønsket å gå i dybden av hvordan definisjoner benyttes i offentlig sektor med relevante eksempler. Resultatet av dette har blitt et i overkant omfattende kapittel. Vi velger imidlertid å beholde dette her for å gi innblikk i ulike sider av definisjoners anvendelser.

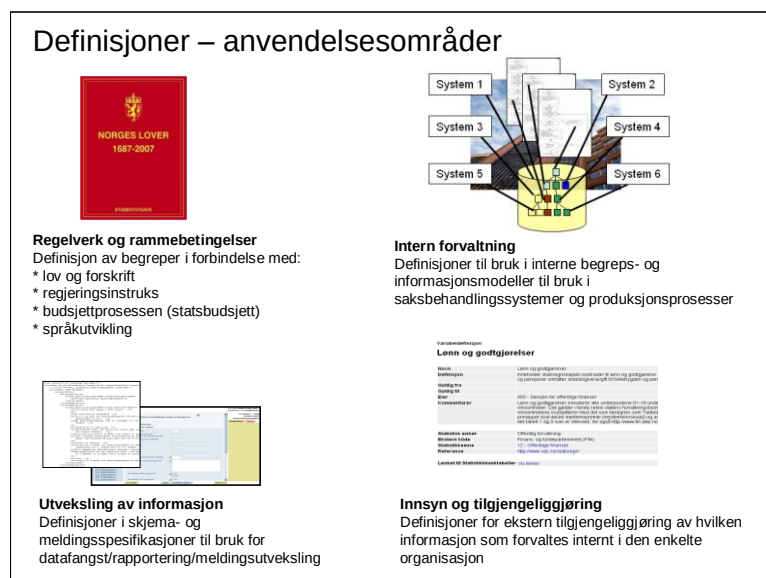
Utgangspunktet for strukturering og systematisering av behovsanalysen er basert på arbeidet med rapporten ”Innspill til metadatastrategi for offentlig sektor”, som ble levert til Fornyings- og administrasjonsdepartementet 10.juni 2009⁴.

Denne rapporten listet opp 4 hovedområder for anvendelser av metadata. I prinsippet er metadata og definisjoner svært nært beslektet. På mange måter kan vi si at en definisjon er selve grunnstammen i hva som definerer et informasjonselement. Metadata favner også de aspektene som eksplisitt uttrykker det faktiske innholdet i et informasjonselement,

⁴ ”Innspill til metadatastrategi for offentlig sektor”, Fornyings- og administrasjonsdepartementet (2009)

selve representasjonen av data. Således vil det være relevant å betrakte anvendelsesområdene for definisjoner på samme måte som for metadata.

Noe revidert ser vi at modellen over anvendelsesområder for metadata kan fungere som en overbygning også over de ulike anvendelsesområder vi kan se for oss for definisjonsarbeid⁵:



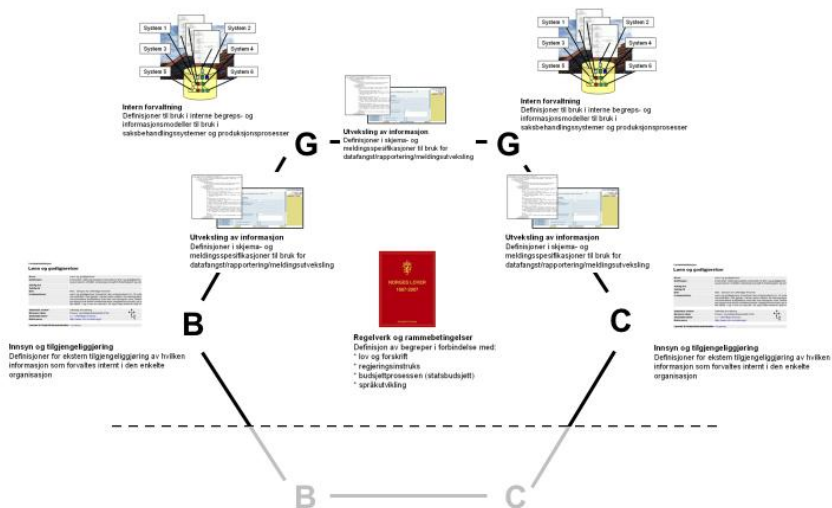
Figur 6 - Overordnet skisse over anvendelsesområder for definisjoner

Gjennom arbeidet med å finne eksempler på anvendelser har vi også vurdert eksempler som enten inngår i de enkelte hovedområdene over, eller er varianter som også refererer på tvers av modellen over de 4 hovedområdene ovenfor. Se beskrivelser av de ulike områdene med eksempler i det følgende.

Figur 7 viser er den enkelte offentlige virksomhet (G) beskrevet ved interne begreps- og informasjonsmodeller, hvor definisjoner vil være viktige referanser i disse. Tilsvarende vil definisjoner også inngå i dialogen mellom offentlige virksomheter (G2G) og mellom offentlige virksomheter og næringsliv (B2G) / innbyggere (C2G)

Samhandling mellom næringsliv og innbyggere (C2C, B2B og B2C) er utenfor scope av denne behovsanalysen siden vi forholder oss til definisjonsarbeid i offentlig sektor.

⁵ Merk at notasjonen for angivelse av områder er endret fra modellen i rapporten "Innspill til metadatastrategi for offentlig sektor" (FAD, 2009)



Figur 7 – Anvendelsesområder i en samhandlingskontekst

3.5.1 Regelverk og rammebetingelser

Regelverk og rammebetingelser favner flere perspektiver og vi har i det følgende sett på definisjoner i forbindelse med forarbeider, lov og forskrift, i forbindelse med det offentlige budsjettprosesser og den generelle språkutviklingen i samfunnet, som alle er opprinnelsessteder førende for all videre bruk av definisjoner.

Disse reglene og rammebetingelsene danner dermed en basis som definisjoner i offentlig sektor må forholde seg til.



Figur 8 - Faksimile fra Aftenposten 8. desember 2009 hvor det etterlyses en endring i definisjonen av fattigdom.

3.5.1.1 Definisjoner i lov og forskrift

I St.Meld. 19 (2008-2009) står det

"Omgrep i lovgivinga bør vere felles og ikkje stå i vegen for elektronisk samarbeid."

I den enkelte etat er lover og forskrifter med forarbeider en viktig del av det rettslige grunnlaget for forvaltningens virksomhet. Lover og forskrifter som påvirker en etats forvaltning danner dermed en ramme for forståelsen av begreper i etaten. Mens det f.eks. i EU er populært at sentrale begreper i en lov eller forskrift er formaldefinert, dvs. definisjonen av dem kommer først i lovtekst (også kalt legaldefinisjoner), har dette ikke vært like vanlig i Norge. EØS-avtalen gjør imidlertid at mange definisjoner kommer inn i norsk rett via direktiver og forordninger fra EU.

En typisk samhandlingsproblemstilling oppstår når en etat, som ikke har begrepet presist definert (kanskje fordi det ikke er sentralt i etatens forvaltningsområde),

Eksempel på inkonsistens i definisjoner på tvers av forvaltningsområder:

Lov om fri rettshjelp sier "Søknad om fritt rettsråd kan innvilges til den som har inntekt og formue under bestemte grenser fastsatt av departementet..."

I Forskrift til lov om fri rettshjelp sies det "§ 1-1. Inntektsgrensen for fri rettshjelp er kr 246 000,- for enslige og kr 369 000,- for ektefeller og andre som lever sammen med felles økonomi."

Det er ønskelig bl.a. utifra å unngå dobbeltrapportering at inntekten hentes fra Skatteetaten, dvs selvangivelsen fra foregående år. Skattedirektoratet har imidlertid 13 definisjoner av bruttoinntekt. Hvilken skal de velge utifra at lovgiver i paragrafene over ikke har angitt presis definisjon? Kunne dette ha vært unngått dersom lovgiver hadde spesifisert klarere?

skal samhandle med en etat som har presise definisjoner.

Hensyn som må sikres i forbindelse med utarbeidelse av definisjoner til bruk i forbindelse med lov- og forskriftsarbeid:

- Definisjoner med bakgrunn i lovtekst må forstås i kontekst av lovområde.
- Det må kunne refereres til alle relevante rettskilder.

3.5.1.2 Definisjonsarbeid i forbindelse med budsjettprosessen

I regjeringens utarbeidelse av statsbudsjett medfører bl.a. endringer i avgiftspolitikken Dette kan skape behov for nye begrep og begrepsdefinisjoner.

Hensyn som bør støttes for utarbeidelse av definisjoner i forbindelse med budsjettprosessen:

- Definisjoner må kunne uttrykkes formelt og få en formell forankring, slik at disse blir konsekvent benyttet i for eksempel videre lov- og forskriftsarbeid.
- Det må være mulig å håndtere endringer i definisjoner over tid og samtidig bevare historikk.

3.5.1.3 Automatisering av lovverk gjennom utvikling av saksbehandlingssystemer

Enkelte forvaltningsområder har et mer automatiseringsvennlig lovverk enn andre⁶. Dette kan ofte ha sammenheng med hvor statisk selve lovområdet er. Det er slik at for områder med stor grad av dynamikk så vil forvaltningen av høy presisjon i lovgivningen og beskrivelsen av de ulike definisjonene ha en høy kostnad, idet definisjonene endres raskt over tid. For mer statiske områder vil det kunne være mer hensiktsmessig med en høyere presisjonsgrad.

For lovområder som automatiseres gjennom systemstøtte har man hatt følgende sekvensielle prosess for utvikling av lov/forskrift og videre systemutvikling

Eksempel på endring i en definisjon i forbindelse med budsjettprosessen:

Biodiesel som definert i wikipedia er drivstoff med vegetabilsk eller animalsk kilde. I Norge produseres biodiesel fremstilt av lakseolje, fiskeavfall, fritryllje og raps. Det er imidlertid kun tilsetninger basert på rapsolje, eller hovedsakelig rapsolje, som tilfredsstiller den europeiske CEN-standard EN 14214.

I statsbudsjettet for 2010 besluttet regjeringen avgift på biodiesel argumentert ut i fra at rapsolje som råmateriale fortrenger matproduksjon. Problemet, og den langvarige politiske diskusjonen, kunne kanskje vært unngått ved å definere biodiesel (i teknisk kontekst) som drivstoff basert på kun de tre første.

Dersom biodiesel var definert som en avgiftsfri diesel, vil definisjonen etter vedtaket være endret (i skatt- og avgiftsmessig kontekst).

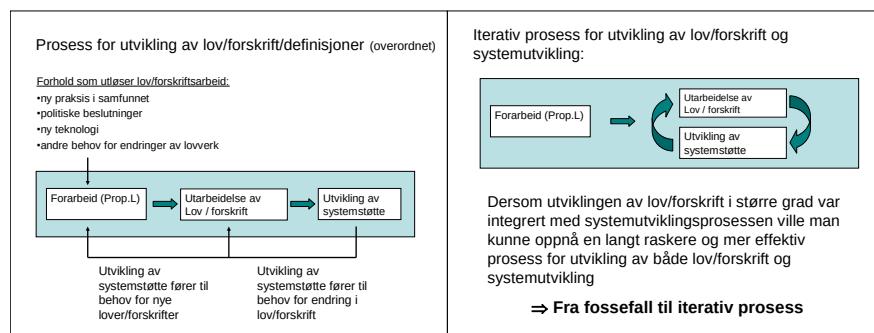
Eksempel på lovgivers systemstøtte

Begrepet "urimelige avtalevilkår" i Avtaleloven er bevisst beskrevet vagt for at rettspraksis skal kunne vurdere hva som til en hver tid betraktes som urimelig.

Det som var urimelig avtale i 1960 er ikke nødvendigvis det i 2009. Men hvor er det beskrevet hva som skal forstås presist, eller som rettspraksis? Burde ikke denne definisjonen beskrive denne bevisste vagheten?

⁶ "Automatiseringsvennlig lovgivning", Jon Bing, 1977

(Figur 9 til venstre nedenfor). Figur 9 til høyre viser hvordan prosessen kan gjøres mer effektiv med bedret systemstøtte.



Figur 9 - Prosess for utvikling av lov/forskrift og systemutvikling. Til venstre: Slik prosessen som regel foregår i dag. Til høyre: Slik prosessen kan effektiviseres.

Hensikten bak lov eller forskrift kan legge føringer for evnen til automatisering. Dersom det er et ønske at rettspraksis skal kunne bestemme definisjonen vil den være vag (f.eks. "urimelige avtalevilkår", tilsvarende om intensjonen er at saksbehandler skal kunne utøve skjønn. Dersom ønsket er en automatisering av saksbehandlingen kan (og bør) definisjonene skrives slik at de er automatiseringsvennlige,

Hensyn som må sikres i ved utarbeidelse av definisjoner i forbindelse med automatiseringsprosess og systemutvikling:

- Støtte ulik grad av presisjon.
- Tettere integrasjon mellom språklig definisjon og IT-teknisk representasjon for systemutvikling.

3.5.1.4 Endringer med bakgrunn i generell utvikling av språk og kulturpåvirkning

Over tid utvikles nye begreper og definisjoner gjennom at språket endrer seg. Dette er ofte drevet av en generell samfunnsutvikling, ofte drevet av nye teknologier. Innføring av e-post er også et godt eksempel på hvordan nye begreper har oppstått i språket, og videre blitt definert inn i lover og forskrifter. Ord og uttrykk kan også bli politisk ukorrekte (f.eks. handikappet) og

Eksempel på tolkning og skjønn

Myndighetsalderen i Norge er 18 år. Med myndighetsalder følger en rekke rettigheter og plikter.

Hva skjer dersom man blir arrestert på 18 års dag? Får man kjøpe alkohol på 18 års dag? Eller, du skader tennene dine, kan du få gratis tannlege? Kan du delta i stortingsvalg?

Det ikke alltid de samme reglene gjelder i forhold til når rettigheter og plikter ved myndighetsalder tar til. "Det året du blir 18" sier valgloven, "den dagen du fyller 18." sier alkoholloven"

Kan tolkningen bli feil fordi den gjøres for langt fra de som utformet lovens intensjon?

Merknad [A1]: Side: 16
Men neppe særlig aktuelt å la inngå i automatiske rutiner.

HAR IKKE ENDRET, FORESLÅR Å BEHOLDE TEKSTEN (LASSE)

Eksempel på endring i en definisjon med bakgrunn i språkendringer

Fra Språknytt 4/2009, Språkrådet: "Få nordmenn snakkar om blackboard, men likevel bruker stadig fleire ordet whiteboard, eller til og med det pleonastiske whiteboard-tavle om kvite tusjtafler. Avlørsarordet tusjtafle har alt fest seg nokolunde som namn på denne tavla – eit like godt motstykke til krittavle som blackboard er til whiteboard."

erstatninger kan ofte medføre små endringer i betydning som vil omfatte smalere eller videre grupperinger.

Hensyn som bør støttes for utarbeidelse av definisjoner i forbindelse med budsjettprosessen:

- Definisjoner må ta hensyn til gjeldende normer gitt av språk og kultur, formidlet av institusjoner som Språkrådet.

3.5.2 Intern forvaltning

Den enkelte forvaltningsenhet er forankret i loven og bygger på definisjoner som illustrert over. Saksbehandling, prosesser og systemer forholder seg til disse rammene og trenger felles definisjoner og fortolkninger.

3.5.2.1 Generell støtte til intern informasjonsforvaltning

I forbindelse med etatens interne saksbehandling og informasjonsforvaltning er det behov for eksplisitte og gode beskrivelser av informasjon og prosesser som dokumenterer og systematiserer etatenes virksomhet. Spesielt er det behov for presise beskrivelser av hvordan informasjon representeres.

Informasjon om begrepene som benyttes finner vi i informasjonsmodeller, prosjektinformasjon mv, men denne informasjonen er normalt lite tilgjengelig i form av felles begreper og begrepsdefinisjoner.

Hensyn som må sikres ved etatenes interne dokumentasjon av informasjon og forvaltning av system og prosesser:

- Gode og eksplisitte beskrivelser av informasjon, systemer og prosesser.
- Tydelig og presis beskrivelse av hvordan informasjonen representeres.

3.5.2.2 Kalkulatorer og simulatorer

I St.Meld. 19. (2008-2009) er det et større avsnitt om rettssikkerhet og automatiserte avgjørelser. I meldingen sies det:

"Dette kan gjere det enklare for innbyggjarane å kontrollere at avgjerdene er rette. Simuleringsprogramma må då vere samanfallande med dei programma som vert brukte i sakshandsaminga." [...] "Regjeringa vil vurdere om dagens plikt til grunngiving kan

Eksempel på simulatorer. Situasjonsbetingete tjenester

Peder Ås og Marte Kirkerud har nettopp blitt foreldre. Sykehuset har sendt varsel til folkeregisteret om den nye innbyggeren, og om et par uker vil Marte få et brev fra NAV som forteller at de har fattet vedtak om tildeling av barnetrygd, kr 970,- per måned fra og med neste måned. Fordi Marte ikke har jobbet minst seks av de siste ti månedene før fødselen, har de ikke krav på "foreldrepenger ved fødsel". Men Peder og Marte har undersøkt på nettsidene til NAV og kommunen de bor i, samt på Minside, og funnet ut at de kan få engangsstønad, som utgjør kr 33.358,-.

Det de ikke vet er at de i den samme perioden også har rett til bostøtte, kr 2.057,- per måned, samt rentefritak på studielånet – en ordning som gir kostnadsfri utsettelse av avdragene. Til sammen ville dette gi dem en økning i disponibel inntekt på nesten 25 %, nærmere bestemt kr 24.324,-, i det første halvåret etter fødselen.

utvidast slik at det vert sett klarare krav om at simuleringsprogram og dokumentasjon av avgjerdssystema vert tilgjengelege.” [...] “Lover og regelverk bør vere utforma slik at dei legg til rette for automatiserte avgjerdssystemar der det er grunnlag for det”

Løsningen er altså at klare og publiserte lover og regelverk hvor definisjoner er utformet slik at simuleringsprogram kan lages og tilgjengeliggjøres av tredjepart.

Hensyn som må sikres ved utarbeidelse av definisjoner for automatiske avgjørelser:

- Entydig og klare definisjoner

3.5.3 Utdveksling av informasjon

Utdveksling av informasjon handler om kommunikasjonen mellom etater, næringsliv/innbyggere og det offentlige. Vi konsentrerer oss om den elektroniske kommunikasjonen og betrakter utdveksling av data på ulike måter; både planlagt og ikke planlagt meldingsutdveksling (skjema og rapportering) samt forhåndsdefinerte rapporteringsordninger og de mer tilfeldige og ikke forhåndsdefinerte utdvekslinger.



Figure 10 - Faksimile fra Aftenposten 2. november 2009. Er dette mer vanlig enn uvanlig? Opererer etatene med ulike definisjoner og regler for ”troverdig ekteskap”?

3.5.3.1 Avtalebasert meldingsutveksling

For at den enkelte etat skal kunne utføre sin forvaltningsoppgave er det behov for informasjon fra andre etater, innbyggere og næringsliv. Dette kan dreie seg om rapportering av inntekts- og formuesstørrelser for at Skatteetaten skal kunne utligne skatt (selvangivelsen), sykemeldinger eller arbeidsledighetsrapporter (meldekort) for beregning og utbetaling av sosiale stønader fra NAV. Det kan også gjelde søknader om ulike tillatelser for drift av næringsvirksomhet, for eksempel søknad om skjenkebevilling, eller innmeldinger eller endringer av opplysninger i ulike registre, for eksempel melding til Folkeregisteret om flytting.

Til disse ulike formål er det utarbeidet egne skjema og rapporteringsordninger for både næringsliv og privatpersoner. Mens de i utgangspunktet ofte har bakgrunn fra fysiske blanketter, blir skjemaene og tjenestene i større og større grad tilbudt elektronisk. Det utarbeides elektroniske skjema og meldingsspesifikasjoner, som definerer alle informasjonselementene som inngår i en slik rapporteringsordning. Til dette formål må det være klare og eksplisitte definisjoner, som både gir en presis beskrivelse av hvilken informasjon som skal rapporteres, og hvilke representasjonsformater som legges til grunn for det enkelte informasjonselement.

Skal etatene utveksle informasjon med hverandre eller næringsliv og personer, må det være enighet i hva som utveksles, ikke bare feltnavn og formater, men også hvordan innholdet skal forstås. Til dette trengs det definisjoner på begrepene feltene representerer. Normalt skjer dette gjennom at de utvekslende partene på forhånd enes om meldingsspesifikasjoner.

Det blir fort kostbart å lage meldingsformater mellom alle aktører⁷ på tradisjonelt punkt-til-punkt transformasjoner. En samordning på meldingsutvekslingsområdet er drevet frem både av intern effektivisering, men også av samhandlingsinitiativer og EU-direktiver og implementert gjennom initiativer som Norge.no og Altinn.

Hensyn som må sikres ved definisjon av skjema- og meldingsutvekslingsformater:

- Gode og eksplisitte beskrivelser av informasjon og dennes kontekst
- Tydelig og presis beskrivelse av hvordan informasjonen representeres.
- Klar kobling datadefinisjoner og begrepene de representerer

Eksempel på avtalebasert meldingsutveksling

Husbanken innførte i 1972 automatisert handsaming av søknader om bustønad (jf. Bing 1984). Skjønnsreglar vart erstatta med klare inntektsgrenser baserte på tal frå skatteetaten. Fødselsnummer vart mata inn i systemet, opplysningar om inntekt vart henta automatisk frå skatteetaten, og informasjon om trygdestatus kom frå trygdeadministrasjonen. Faktiske og rimelege utgifter vart berekna på grunnlag av dei opplysningane banken sjølv hadde om bustaden. Programmet avgjorde så om søkjaren hadde rett til bustønad. I så fall vart ein utbetalingsgiro skriven ut.

Fra St.meld.19 (2008-2009)

⁷ Ettersom antall aktører øker vil en punkt-til-punkt meldingsutveksling kreve $n*(n-1)$ transformasjoner. Ved å basere seg på felles modeller vil den reduseres til n transformasjoner

3.5.3.2 Sammenstilling av innhold, åpne data og ad hoc samhandling

Definisjoner må være gode nok til at aktører kan ta avgjørelser omkring nytten av og risikoen ved å bruke data fra andre kilder enn de selv kontrollerer. Dette kan innebære at man må kunne beskrive forutsetningene og konteksten for avgivning av dataene (slik som tidspunkt, periode, hvem som har avgitt dataene, i hvilken sammenheng), hvilket går utover det terminologiske.

I motsetning til avtalebasert meldingsutveksling er det ofte tilfeller hvor man har behov for informasjon fra kilder hvor man ikke har anledning til å enes om formater og semantikk. Det er derfor behov for klare semantiske beskrivelser og av innhold.

Hensyn som må sikres ved definisjon av åpne data og sammenstilling:

- Behov for eksterne semantiske definisjoner som beskriver begreper brukt i en utveksling

Eksempel på sammenstilling av innhold fra åpne datakilder

Byantikvaren ønsker å se sammenhengen mellom avslag på byggesaker og branner for verneverdige bygninger. Utgangspunktet er: Byggesaker er tilgjengelig hos kommunen, Byantikvarens liste over verneverdige bygninger – den såkalte gule listen - og brannvesenets oversikt over bygningsbranner i Oslo. Gjennom å sammenstille denne informasjonen finner han en påfallende sammenheng i noen av tilfellene.

3.5.4 Innsyn og tilgjengeliggjøring

Innsyn og tilgjengeliggjøring handler om å ivareta den enkeltes krav til innsyn i hvordan forvaltningen opererer på vegne av samfunnet. Det er en utfordring for forvaltningen operere så transparent at alle kan forstå hvordan de ulike forvaltningsområdene fungerer.

En av utfordringene er å tilgjengeliggjøre hvilken informasjon det offentlige forvalter. Det må være mulig for alle å få tilgang til begreper og definisjoner på en hensiktsmessig måte.

Tilgjengeliggjøring dreier seg også om å tilgjengeliggjøre tjenester og informasjon, og det å utnytte definisjoner til gjenfinning.

3.5.4.1 Gjenfinning

Det må være mulig å benytte begrepene og strukturene til å finne informasjon og tjenester i de ulike etatene. Søk på for eksempel på inntektsbegreper burde resultere i ressurser (tjenester, dokumenter osv.) relatert til begrepet og relasjoner til andre relevante begreper og deres ressurser.

Hovedhensyn som må sikres ved definisjon brukt i en

Eksempel på gjenfinning

På statens vegvesen sine sider kan man lese

"I 1995 ble ansvaret for driften av vegnettet organisatorisk slått sammen med Biltilsynets tjenester og andre trafikkfunksjoner i en ny trafikkavdeling. Etter dette ble både "Driftsavdelingen" og "Biltilsynet" borte som betegnelser på enheter i Statens vegvesen. Utestasjonene fikk samtidig betegnelsen trafikkstasjoner."

På tross av dette har "Biltilsynet" satt seg som navn i allmennheten, og bør derfor kunne brukes som en uoffisiell term i gjenfinningsøyemed.

gjenfinnings- og søkekontekst:

- Gode og eksplisitte beskrivelser av informasjon og dennes kontekst
- Referanser og relasjoner til andre og liknende definisjoner
- Eierskap til definisjonen
- Representasjon må kunne beskrive relasjoner mellom begreper

3.5.4.2 Publisering av begrepsdefinisjoner

For å sikre god oversikt over hvilke begrepsdefinisjoner en etat forvalter bør disse tilgjengeliggjøres med felles teknikker slik at omverdenen lett kan få tilgang til dem. Både generelle definisjoner, kontekst og relasjoner til andre begreper og forvaltningsområder vil være nyttige og relevante elementer i beskrivelsene av den enkelte definisjon. Her er det mye å hente fra etater og initiativer som Statistisk Sentralbyrå, KITH og Brønnøysundregistrene som allerede har tilgjengeliggjort definisjonene på nett.

Hovedhensyn som må sikres ved definisjon brukt i en gjenfinnings- og søkekontekst:

- Gode og eksplisitte beskrivelser av begreper og begrepsdefinisjoner og dennes kontekst
- Referanser og relasjoner til andre og liknende definisjoner
- Eierskap til definisjonen
- Må kunne eksportere begreper og definisjoner lesbare for både mennesker og maskiner.

Eksempel på publisering av begrepsdefinisjoner

I SERES blir alle definisjoner gjort tilgjengelige som etatens egen ordbok. Definisjonene får en unik identifikator (URL) som viser dets forvalter/eier, kontekst og preferert term. Ved å slå opp med URLen returneres både definisjonen i HTML, men også maskinlesbare formater som RDF.

Med begrepsdefinisjoner publisert som dette vil man kunne benytte definisjonene som rene ordbøker, som annoteringer i meldingsutveksling og åpne data, og som grunnlag for harmonisering.

4 Kriterier

4.1 Krav til for definisjonsarbeid i offentlig sektor

I kapittel 3.5 har vi kartlagt områder hvor mangel på gode begrepsdefinisjoner skaper utfordringer, og hva som kreves av en god begrepsdefinisjon i disse tilfellene. I det følgende gjennomgår vi noen av disse kriteriene samlet. Kravene under kan sees på som utdyping av Standardiseringsrådets kriterium 4. Se bilag 2 for en detaljert gjennomgang av hver enkelt standard mot kriteriene.

4.1.1 Allmenn forståelse og språk

Begrepsdefinisjonen bør som hovedregel bryte minst mulig med *normal forståelse av ordet eller uttrykket*.

Et annet viktig aspekt er muligheten for å angi *synonymer* (alternative termer) for begrepet. Videre kan det også være av interesse å kunne angi ikke-prefererte termer – f.eks. ord som kan være ønskelig å benytte i søk, men ikke vises.

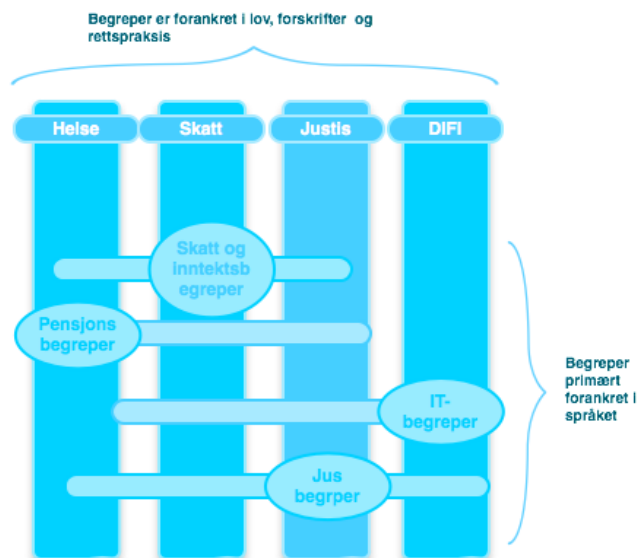
Begrepsdefinisjon og term er alltid uttrykt i et *språk*. Det bør det være mulig å uttrykke både *term og begrepsdefinisjon i flere språk*. I offentlig sektor er dette viktig ettersom vi har flere offisielle språk (norsk, samisk). Det er imidlertid store farer ved å gjøre dette ukritisk ettersom betydningen fort kan bli forskjellig fra et språk til et annet.

4.1.2 Kontekst

Ord og uttrykk endrer mening fra én sammenheng til en annen. Begrepet må derfor forstås innenfor *sin sammenheng*. Mens begrepet er ”det man snakker om”, er konteksten ”sammenhengen det skal forstås”. Termen kan vise til ulike begrep avhengig av kontekst er boligmarkedet eller pokerspill.

I offentlig sektor er etatens fagspråk gitt av etatens fagområde, og beskrevet i rettskilder som lov- og forskrift. I sin kontekst har etatens behov for *adskillende egenskaper* (som for utenforstående ikke nødvendigvis ser relevant ut). *Eierskapet* til slike begreper og definisjoner ligger i den enkelte etat, og deres bruk av begrepet. Når en etat benytter begreper utenfor sin etats fagområde, forankret i annen rettspraksis eller lekmannsspråk, bør de kunne *referere* til denne etatens definisjoner.

Semantiske misforståelser oppstår gjerne når man tror man operere innen samme kontekst, men det er forskjell i detaljnivå fagområdene i mellom. Man trenger derfor også mulighet for å uttrykke kontekst.



Figur 11 - Begreper sentralt innen en etats fagområde vil helst være forankret i lov, forskrifter eller rettspraksis – og dermed tilhøre denne etaten og ha en detaljrikdom faget krever. Begreper som ikke er så sentralt fagområdet til etaten, vil ofte være mer definert i andre etater eller lekmannspråk og bør derfor relatere seg til disse.

4.1.3 Definisjonens innhold, nøyaktighet og kvalitet

Gode definisjoner bør skrives etter bestemte regler eller metodikk. De bør være korte og beskrivelsen må være forståelig for *målgruppen*.

Det skilles gjerne mellom innholdsdefinisjoner (engelsk: intensional), som fokuserer på *adskillende egenskaper*, og omfangsdefinisjoner (engelsk: extensional), som lister referentene som inngår i begrepet. Selv om definisjonene bør kunne skrives på begge vis, bør de som hovedregel bør definisjoner initielt skrives med utgangspunkt i adskillende kjennetegn. Dette vil være et godt utgangspunkt for strukturering og automatisering.

F.eks. definisjon av Fylke bør både kunne beskrives som en første-nivå inndeling av Norge, og som en liste av de fylker som inngår. Korrekt bruk av innholdsdefinisjoner og omfangsdefinisjoner sier noe om nøyaktigheten til definisjonen.

Andre *kvalitetsmessige sider* ved en definisjon fravær av sirkelreferanser og negative definisjoner.

4.1.4 Strukturering av begreper

Mange kunnskapsorganiseringsapparater, slik som taksonomier, thesaruri, klassifikasjonsskjema, emnekart, deler en felles struktur, og brukes i liknende anvendelser. Relasjoner mellom begrep kan deles inn i grupper

- Generiske relasjoner (en traktor *er* et kjøretøy)
- Partitive relasjoner (et hjul kan være *del av* en traktor)
- Assosiative relasjoner (en gård *driver med* korn)
- Relasjoner som bygges mellom begrepsystemer (en <turisme>hytte *tilsvarer* en <skatt>fritidseiendom)

Det bør være beskrevet metodisk hvordan struktureringen gjøres.

4.1.5 Refererbarhet

Vi anser at målet med refererbarhet er å kunne referere unikt til et begrep (med dets kontekst i form av rettskilder m.m.) og benytte denne referansen til å returnere en definisjon. Aktuelle representasjonsstandarder bør benytte URIer⁸ i form av URLer⁹ til dette formålet hvor identifikatoren inneholder både *begrepets eier eller forvalter* og *begrepets prefererte navn*. URLer anbefales over URIer fordi de kan returnere den faktiske definisjonen og andre metadata til begrepet.

Definisjoner bør være tilgjengelig for alle og forvaltes som et helhetlig kodeverk av etaten. På denne måten kan definisjonene bli harmonisert og gjelde for alle deler av virksomheten. Dette vil være et viktig bidrag til å få et begrepsapparat som er avstemt på tvers av prosjekter, fagområder og organisasjonsenheter.

Our public data principles state that:

- Public data will be published in reusable, machine readable form
- Public data will be available and easy to find through a single easy to use online access point (www.data.gov.uk)
- Public data will be published using open standards and following the recommendations of the World Wide Web Consortium
- Any 'raw' dataset will be represented in linked data form
- More public data will be released under an open licence which enables free reuse, including commercial reuse
- Data underlying the Government's own websites will be published in reusable form for others to use
- Personal, classified, commercially sensitive and third-party data will continue to be protected."

"Government public data principles", UK government desember 2009

⁸ Uniform Resource Identifier er en tegnstring som representerer en Internett-ressurs, f.eks. <mailto:ola.normann@norge.no> eller <http://www.difi.no>. Generelt finnes mange systemer for maskinlesbare referanser, identifikasjonsskjema som ISBN, UUID, GUID, OID, URI m.fl. Med suksessen til World Wide Web på 90-tallet, satte W3C og IETF en arbeidsgruppe som resulterte i RFC 3305 - "Report from the Joint W3C/IETF URI Planning Interest Group: Uniform Resource Identifiers (URIs), URLs, and Uniform Resource Names (URNs): Clarifications and Recommendations".

⁹ Uniform Resource Locator er en tegnstring som representerer en side på World Wide Web, f.eks. <http://www.difi.no>. URL'er er et subset av URIer. URIer oppfyller også muligheten til å returnere en metadata beskrivelse av begrepet og dets definisjon, ikke bare selve ressursen.

4.1.6 Endringshåndtering og sporbarhet

Endringer er et faktum, også for begreper og begrepsapparater. Endringer i begreper kan skje gradvis, men kan også være større og skje på et bestemt tidspunkt f.eks. gjennom forskrift. Det er viktig at denne dynamikken støttes i standardene. Som et minimum bør en standard adressere metodikk som sier noe om hva som kvalifiserer som til en endring, sletting av begreper, gyldighet og stabilitet. Standarder bør også angi sporbarheten ivaretas.

Til høyre vises et eksempel på slike regler.

4.1.7 Koordinering - Gjenbrukbarhet og harmonisering

Harmonisering og gjenbrukbarhet kan sees i sammenheng. Der gjenbrukbarhet betyr at man tar utgangspunkt i et begrep og forfiner det, har harmonisering som mål at man skaper et felles begrep.

Et *desentralisert eierskap* vil muliggjøre at man både kan referere og raffinere definisjoner beskrevet innenfor og utenfor det aktuelle begrepsapparatet.

Harmoniseringsstandarter reflekterer gjerne en styrt prosess. Det kan være verdtfullt å se om man kan benytte mer *konvergerende metoder* (Web 2.0) for å oppnå harmonisering. Desentralisert eierskap kan også innebære at aktørene kan arbeide uavhengig av hverandre, noe som gjør samordning mer dynamisk.

Postulat: Man lager et eget begrep av to årsaker; enten fordi man faktisk har behov for en distinksjon, eller fordi man ikke finner et hertil egnet begrep. Tilgang til begreper og gode definisjoner vil kunne skape gjenbruk og dermed harmonisering ved "latskap" fordi det blir lettere å gjenbruke.

4.2 Presiseringer og avgrensninger

Det har vært nødvendig å gjøre noen avgrensninger i forbindelse med hvilke standarder som skal inkluderes i denne vurderingen. I motsetning til tidligere standarder som har blitt vurdert av Standardiseringsrådet er det her mange standarder som har sin utbredelse innenfor en bestemt sektor. Vi har valgt å gjøre noen presiseringer her som avgrenser standarder som blir vurdert.

Et begrep skal oppfattes som uendret når;

- *Det legges til en begrepsdefinisjon der det ikke var noen fra tidligere.*
- *Begrepet ender administrativ status til deprecated (eller slettes)*

Et begrep skal oppfattes som endret når

- *Navnet på konseptet (termen) endres*
- *Definisjonen endres semantisk (meningsinnholdet er endret)*

Dersom to begreper "Elektrisk motorsykkel" kan oppfattes som et underbegrep av et videre begrep "Elektriske Kjøretøy", vil man kunne kommunisere én vei, data kan flyte fra den spesielle til den generelle, og en ytterligere harmonisering kan være unødvendig.

Et begrep "Fritidseiendom" er et presist – men vidt - begrep i skattesammenheng som bl.a. bygger på kravet om bruk til fritidsformål, og kan dermed omfatte byleilighet. Begrepet "Hytte" som f.eks. benyttes i dagligtale kan omfatte, skattemessig en bolig dersom den beboes hele året. Berepene er altså ikke mulig plasseres hierarkisk, og harmonisering kan være ønskelig for å kunne samhandle.

Tabellen under oppsummerer viktigheten til kriteriene innenfor begrepsdefinisjonenes ulike anvendelsesområder.

Følgende notasjon er brukt i tabellen

- = ikke viktig
- + = viktig
- ++ = meget viktig
- +++ = helt nødvendig

Tabell 3 - Kriterier vurdert opp mot anslått viktighet i anvendelsesområdene.

	Regelverk og ramme-betingelser	Intern forvaltning	Utveksling av informasjon	Innsyn og tilgjengelig gjøring
Allmenn forståelse og språk	++	+	-	++
Kontekst	++	+	++	++
Definisjoners innhold, nøyaktighet og kvalitet	++	++	++	+
Strukturering av begreper	-	++	++	+
Refererbarhet	++	+	+	+
Endringshåndtering og sporbarhet	+	++	++	+
Gjenbrukbarhet og harmonisering	-	+	++	+

5 Vurdering av standarder

Kapittel 2.1 legger premissene for vurdering av aktuelle forvaltningsstandarder. Disse er lagt til grunn ved utvalg av standardene.

5.1 Avgrensninger

Mengden av standarder som *berører* dette området er stort, og det er derfor behov for å avgrense til de som har definisjonsarbeid som hovedfokus. Gjennom å studere standardene oppdages også referanser til andre standarder, noen av disse er utelatt ettersom de ikke oppfattes å ha særlig utbredelse.

I samråd med Difi har arbeidsgruppen har gjort noen ytterligere avgrensninger:

1. Internasjonale standarder som ikke er fullført sitt løp gjennom standardiseringsorganisasjonene, eller deler av standarder som ikke har fullført sitt løp vil ikke vurderes.
2. Internasjonale initiativer som ikke er underlagt et internasjonalt standardiseringsorganisasjon vil ikke bli vurdert.
3. Standarder som oppfattes som underdokumenterte blir ikke vurdert.
4. Standarder som har sitt nedslagsfelt i et bestemt domene; dvs. er i anvendelse kun i et domene – blir ikke vurdert. Med domene menes her f.eks. forsvar, helse, finans osv.
5. Standarder som er generiske og ikke selv har hovedfokus på opprettelse av eget begrepsapparatet (semantikk).

Nasjonale initiativer vil bli vurdert om de dekker flere domener, har tilstrekkelig dokumentasjon og har et løp mot standardisering.

Endel standarder som utenfra kan sees som relevante, men som vi basert på disse avgrensningene har utelatt i den følgende vurderingen nevnes i det følgende. Vi er klar over at mange er opptatt av noen av disse standardene, og nevner derfor noen av disse her som eksempler og kort begrunnelse for avgrensingen.

OMG UML og OMG OCL

Unified Modeling Language (UML) fra OMG inneholder et sett av grafiske notasjoner, beskrankninger (OCL) og serialiseringer (XMI) som primært benyttes i spesifikasjonsfasen (modellering) i systemutviklingsprosjekter.

I definisjonsprosessen kan UML sees på som et representasjonspråk hvor UML-klasser kan benyttes som en notasjon for begreper. UML har imidlertid ikke et direkte forhold til begreper, begrepsdefinisjoner og relasjoner mellom disse, slik at det må lages en profil (eller veiledning) for UML for støtte i denne prosessen. UML er derfor utelatt i det følgende som en kandidat for forvaltningsstandard i definisjonsarbeid.

Det bør imidlertid nevnes at ettersom Meta Object Facility (MOF)-standarden (meta-metamodell), også fra OMG, får utbredelse og verktøystøtte, vil UML i større grad benyttes som en av flere visualiseringer over MOF-baserte representasjonsspråk som Topic Maps, RDF og OWL.

OMG ODM

Ontology Definition Metamodel (ODM) er en spesifikasjon som definerer en rekke metamodeller og deres samspill. Det lages videre UML profiler som dermed kan gi verktøystøtte for ISO 13250 Topic Maps, RDF og OWL for verktøy som støtter ODM. Fra spesifikasjonen til ODM heter det:

There are three metamodels that represent more structural or descriptive representations that are somewhat less expressive in nature than CL and some DLs.

These include metamodels of the abstract syntax for RDFS, OWL and TM. RDFS, OWL, and TM are commonly used in the semantic web community for describing vocabularies, ontologies, and topics, respectively. Two additional metamodels considered essential to the ODM represent more traditional, software engineering approaches to conceptual modeling: UML2 and ER (Entity Relationship) diagramming.

Vi anser basert på dette at ODM i seg selv ikke er relevant for definisjonsarbeid, men heller underbygger at definisjonsarbeid bør representeres i standardene som nevnt over. Det skal også nevnes at ODM er en ny standard (2009) og forventes utbredelse og verktøystøtte over tid, men ikke kan anses som tilstrekkelig støttet p.t.

W3C XML/XSD

XML Schema er en svært utbredt standard, som benyttes primært i meldingsutveksling. Standarden representerer derimot kun syntaktiske og strukturmessige sider av en medling, og kan derfor ikke vanskelig sees som en standard som fokuserer på definisjon av begreper og begrepsrelasjoner.

OMG BPMN

Business Process Management Notation (BPMN) er en grafisk notasjon for beskrivelse av forretningsprosesser. Notasjonen beskriver prosesser og data-utveksling, men har ingen forhold til begrepsdefinisjoner.

HL7 (3)

Health Level 7 (HL7) er en organisasjon og et rammeverk (sett av standarder) for bl.a. meldingsutveksling i helsesektoren. Meldingsutviklingsstandardene definerer hvordan informasjonens pakkes, språk, struktur og datatyper for sømløs utveksling.

I Norge er HL7 besluttet brukt for sykehustjenesten. Vi forfølger den derimot ikke videre i denne rapporten ettersom det både er en domenespesifikk standard og primært utveksler forhåndsdefinert informasjon (klinisk informasjon) og ikke benyttes til å beskrive nye begreper og begrepsdefinisjoner.

UN/CEFACT Core Components

Core Components (CC) fra UN/CEFACT er syntaksnøytrale og teknologi uavhengige byggeklosser som kan brukes, og brukes, til datamodellering spesielt innenfor e-handel. Hovedformålet er å gjenbruke eksisterende artifakter med felles forståelse som kan brukes på tvers av domener. CC er implementert i meldingsstandardene ebXML og UBL.

CC har fokus på meldingsutveksling, men er ikke ment for definisjonsarbeid der etater selv skal kunne definere sine begreper. CC anses derfor å være utenfor omfanget til denne rapporten.

Vi avgrenser oss her til å komme med en oppfordring om å se på CC i en harmoniseringsprosess. CC er harmoniserte byggeklosser med definert semantikk (f.eks. Adresse) bør CC kunne refereres til av løsninger som f.eks. SERES og matrikkelen slik at det kan genereres meldinger som er CC kompatible.

UN/CEFACT EDIFACT

Electronic Data Interchange For Administration, Commerce and Transport (EDIFACT) er en internasjonal standard utviklet av FN. EDIFACT er også kalt ISO 9735. EDIFACT tilbyr et sett av regler for strukturering av data, en utvekslingsprotokoll og standard meldinger for utveksling. EDIFACT er en pre-XML standard, og offattes i mange miljøer som for lite fleksibel, dyr å ta i bruk (manglende verktøystøtte i forhold til XML) og lite lesbar. Motargumentet er at standarden inneholder en rekke ferdigdefinerte meldingsmodeller og har stor utbredelse innen store industriområder som elektronikkindustrien.

EDIFACT er selvsagt relevant i meldingsutveksling, men oppfattes ikke som relevant for en etats behov for å definere nye meldinger i en samhandling.

5.2 Standardene og anvendelser

Standardene er initielt valgt ut gjennom forespørsler til standardiseringsorganisasjoner og deltakernes kjennskap til nasjonale og internasjonale initiativer. Noen nye standarder har også dukket opp gjennom intervjuer og arbeidsmøter. Standardene er forsøkt plassert i sammenheng av prosessen som vist tidligere. Dette er gjort for å kunne sammenlikne epler og epler og ikke epler og pærer. Tabellen under oppsummerer dette. Cellene angir følgende:

- standarden dekker ikke dette forholdet
- + standarden dekker dette området
- ++ standarden anses som sentral på dette området
- +++ standarden har dette område som formål
- blank ikke sentral del av standarden

Standardene beskrives i mer detalj i 5.4, og detaljer rundt standardene er angitt i vedlegg.

Tabell 4 - Vurderte standarder og deres samsvar med anvendelsesområdene.

	Metode for utvalg av begreper	Metode for strukturering av begreper og definisjons-skriving	Koordinering (Referering og Harmonisering)	Representasjons-språk
ISO 704:2009 "terminologiarbeid – prinsipper og metoder"	+	+++	-	-
ISO 10241:1992 "international terminology standard – preparation and layout"	+	++		
NS ISO 1087:1992 "Terminologiarbeid Ordliste"		++	+	+
ISO 860 "Terminology work – Harmonization of concepts and terms"			++	
NS-EN 45020:1998 "Standardization and Related Activities - General Vocabulary"		+		
Språkrådet - Termlosen	+	++	+	-
NS-ISO 12616:2002 "Oversettelsesrettet terminografi"		+		
ISO 11179-5				+

"Information technology – Specification and standardization of data elements – (5) Naming and identification principles for data elements				
ISO 11179 "Information Technology – Metadata registries"		+		+
ISO 2788:1986 "Guidelines for the establishment and development of monolingual thesauri"	++	+	-	-
Princeton Wordnet	++	+	-	-
Brønnøysundregisterne SERES "Semantikk registeret for elektronisk samhandling"	++	+	+	+
W3C SKOS "Simple Knowledge Organizational System"	-	+	-	++
ISO IEC 13250:2003 Topic Maps	+	+		++
ISO 15926 Integration of Life Cycle data for process plants including oil- and gas production facilities	+	+		+
W3C OWL Web Ontology Language	+	+		+
ISO 15836 Dublin Core Metadata Initiative – Metadata Terms	-	+	-	+

5.3 Om standardiseringsorganisasjoner og initiativene

Standardene er utviklet innenfor et standardiseringsregime gitt av internasjonale standardiseringsorganisasjoner¹⁰ eller nasjonale initiativer. De aktuelle er:

ISO-TC37. ISO (International Organization for Standardization) er verdens største utvikler og publiserer av internasjonale standarder. ISO/TC37 - Grunnlagt i 1936 (som ISA/TC37) under lederskap av østerrikeren Eugen Wüster.

W3C (World Wide Web Consortium) er primære internasjonale standardiseringsorganisasjonen for World Wide Web. W3C ble grunnlagt i 1994 av Tim Berners-Lee som fortsatt leder organisasjonen. Consortiet består av

¹⁰ Se også <http://www.itst.dk/arkitektur-og-standarder/Aabnestandarder/abne-standarder-for-software-i-det-offentlige-1/baggrundsrapporter/Evaluation%20of%20Ten%20Standard%20Setting%20Organizations.pdf>

medlemsorganisasjoner med fulltidsansatte med fokus på å jobbe sammen for å utvikle standarder for WWW. Per september 2009 har organisasjonen 356 medlemmer.

DCMI (The Dublin Core Metadata Initiative) åpen organisasjon opptatt av utvikling av interoperable metadata standarder som støtter et bredt utvalg av bruksområder. Mest sentral er settet av metadata som beskriver ressurser.

Princeton/Global Wordnet Association er en fri, offentlig, ikke kommersiell organisasjon som er en nettverk for å diskutere, dele og samle wordnets for alle verdens språk.

5.4 Kort om de vurderte standardene og initiativene

ISO 704:2009, Terminologiarbeid – Prinsipper og metoder beskriver oppbyggingen av konseptsystemer og deres definisjoner. Har fokus på konsistente og avgrensede begrepssystemer bestående av opptil ca 200 begreper (og relasjoner mellom disse), slik man kan enes om i en ekspertgruppe.

ISO 10241:1992, International terminology standards – Preparation and layout; beskriver ensformig måte å organisere terminologiarbeid og garantere konsistente og koherente begrepssystemer med lesbarhet der ”det ultimate målet for standardiseringsprosessen er å oppnå et normativt vokabular der bare en term relaterer til bare et konsept og omvendt (oversatt fra kap. 4)

NS-ISO 1087-1:2000, Terminologiarbeid – Ordliste – Del 1: Teori og anvendelser gir en ordliste med nyttige engelske og norske begreper for begrepssystemer med deres (engelske) definisjoner, slik de er brukt i ISO 704

NS-ISO 1087-2:2000, Terminologiarbeid – Ordliste – Del 2: Datamaskinelle anvendelser) gir en ordliste relevant for datasystemer som skal støtte terminologiarbeid.

ISO 860:2007, Terminology work – Harmonization of concepts and terms beskriver hvordan begrepssystemer med overlappende beskrivelsesområde kan samordnes, - og situasjoner der man ikke bør samordne. Beskriver harmonisering ved design av en ekspertgruppe, ikke gjennom naturlig konvergens.

Termlosen, Heidi Suonuuti, Termlosen (Språkrådet, 2008). Kort innføring i begrepsanalyse og terminologiarbeid, utgis av språkrådet og er ikke en standard, men gir en ”folkelig og nyttig tilgjengeliggjøring” av ISO 704, 860 og 20241 (se egen beskrivelse). Termlosen, skrevet av tidligere leder av ISO TC-37.

ISO/IEC 11179-5:2004, Information technology – Metadata registries (MDR) beskriver semantiske forhold knyttet til data, hvilket i vårt tilfelle kan være definisjoner. Standarden er mangfoldig ved at

- part 3 beskriver vanlige metadata med bred anvendelighet
- part 4 beskriver regler for definisjonsskriving
- part 5 beskriver identifikasjon og navneprinsipper, og
- part 6 beskriver krav til livssyklusaspekter for administrativ status

NS-ISO 15926, Integration of life-cycle data for process plants including oil and gas production facilities er spesielt tilpasset, - men ikke begrenset til -, begrepssystemer ("referansesystemer") for utstyr i et livssyklusperspektiv der utstyr over tid skal tilfredsstille designkrav i kontekst av et funksjonell kontekst.

NS-ISO 12616:2002, Oversettelsesrettet terminografi eksemplifiserer de terminologiske prinsippene for oversettelse mellom språk, - der forskjeller i kultur og konvensjoner krever spesiell håndtering for å ivareta semantikk ved oversettelsen.

ISO 12620:1999, Computer applications in terminology – Data categories; under revisjon og sterkt utvidelse (Terminology and other language and content resources – Specification of data categories and management of a Data Category Registry for language resources); Beskriver datakategorier (relevante metadata) for "alle typer av språklige ressurser" ("It specifies mechanisms for creating, selecting and maintaining data categories, as well as the interchange format for representing them")

Bør undersøkes ved tilgjengeliggjøring.

W3C SKOS, Simple Knowledge Organization System forsøker å sy sammen en rekke forhold nødvendig for å kommunisere kunnskap, slik som å strukturere den og enkelt publisere den på internett på en slik måte at semantic-web-teknologien kan utnyttes for sammenstilling og tilgjengeliggjøring.

ISO 15836 (Også publisert som ANSI/NISO Z39.85 og IETF RFC5013), Dublin Core Metadata Terms Set angir blant annet 15 vanlige metadata spesielt brukt til å beskrive produksjoner slik som bøker, musikk, osv. De er contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title, type. Inneholder også et utvidet sett.

ISO 13250:2003, Information Technology – SGML applications - Topic Maps er en ISO-standard for representasjon og utveksling av strukturert og semistrukturert informasjon. Standarden startet som en generalisering av konseptene rundt stikkordregister, ordbøker og synonymordlister, men har vist seg å ha stor nytteverdi i en rekke andre sammenhenger. Emnekart brukes til informasjonsgjenfinning og navigering i informasjon.

NS-EN 45020:1998 "Standardization and Related Activities - General Vocabulary". Også kjent som ISO Guide 2.

ISO 2788:1986, Documentation – Guidelines for the establishment and development of monolingual thesauri beskriver metoder for og retningslinjer på linje med ISO 704 for etablering av begrepsapparater, men der denne standarden tar utgangspunkt i riktig valg av termer i stedet for konsepter.

Brønnøysundregisterne SERES, Semantikkregisteret i offentlig sektor. Et norsk initiativ for å understøtte Oppgaveregisteret og Altinn II. Registeret vil inneholde begreper brukt i samhandling for etater som deltar i Altinn samarbeidet.

W3C OWL, Web Ontology Language. OWL er en sentral standard innenfor Semantisk Web. Bygget for å lage formelle begrepsapparater med tanke på å kunne lage logiske slutninger og gjøre konsistenssjekker i begrepsapparatet.

En rekke relevante standarder fra ISO TC37 er under utarbeidelse. Disse er ikke tatt med i vurderingen i det følgende, men kan vurderes ved utgivelse. Dette inkluderer ISO 22274, Systems to manage terminology, knowledge and content – Internationalization and concept-related aspects of classification systems; ISO/TR 24156, Guidelines for using UML notation in terminology work; ISO 24610, Language resource management – Feature structures; ISO 24611, Language resource management – Morpho-syntactic annotation framework; ISO 24612, Language resource management – Linguistic annotation framework (LAF); ISO 24613, Language resource management – Lexical markup framework (LMF); ISO 24615, Language resource management – Syntactic annotation framework; ISO 24617-1, Language resource management – Semantic annotation framework (SemAF) – Part 1: Time and events; ISO 24617-2, Language resource management – Semantic annotation framework (SemAF) – Part 2: Dialogue acts; ISO 26162, Computer applications in terminology – Design, implementation and use of terminology management systems; under utarbeidelse (CD).

5.5 *Evaluerings mot kriteriene*

Tabellen under viser en evaluering av standardene mot kriteriene gitt over. Vi har her brukt følgende vurderingsvekter

- standarden dekker ikke dette kriteriet
- + standarden dekker dette kriteriet
- ++ standarden anses som sentral på dette kriteriet
- +++ standarden har dette kriteriet som formål, og bør anvendes til dette
- blank ikke sentral del av standarden

Table 5 - Sammendrag av vurderinger av standarder opp mot standardiseringsrådets kriterier med detaljering av behovskriterier relatert til området definisjonsarbeid

	Formål med standarden	Proessen rundt forvaltning av standarden prosessen	Standardens aksept og utbredelse i markedet	Behov (sammendrag av de neste kolonnene)	Allmenn forståelse av språk	Kontekst	Definisjoners innhold, nøyaktighet og kvalitet	Strukturering av begreper	Refererbarhet	Endringshåndtering og sporbarhet	Gjenbrukbarhet og harmonisering
ISO 704:2009	+	++	++	+		++		++	-	-	-
NS ISO 1087:2000		++	++								
NS ISO 10241:1987		++	++	+	++						-
ISO 860:2007		++	+	-				+		-	-
NS-EN 45020:1998		++									
Termlosen	+	++	++	++		++		+++	-		
NS-ISO 12616:2002		++									
ISO/IEC 11179-5		++		+					++	+	+
ISO/IEC 11179		++		+	+	++				+	
Wordnet		+	++	++	+	+	+	++	+		
SERES		+	+	++	-	++		+	+	++	+
W3C SKOS		++	++	++		+		++	+		
ISO 13250 Topic Maps		++	++	++		+		++	+		
ISO 15926		++	+	0		+					
W3C OWL		++	++	+		+	+				
ISO 15836 DCMI Metadata Terms		++	+++	+		++					

6 Rapportens anbefalinger og forslag til videre arbeid

Vårt utgangspunkt har vært at krav til definisjonsarbeid avhenger av hva definisjonene skal benyttes til (målgruppe og anvendelse). Uavhengig av bruksområdene kan det likevel trekkes felles slutninger som danner grunnlaget for følgende anbefalinger til Difi og Standardiseringsrådet.

(1a) Det anbefales å basere terminologiarbeidet (metodisk identifisering og strukturering) på Termlosen med refererte ISO standarder ISO704 og ISO1087-1 (og ISO10241).

(1b) Det anbefales å utarbeide norskspråklig versjoner av disse standardene.

(1c) For å ta dette i bruk, anbefales det videre at det utarbeides en veiledning basert på det ovenstående konkret for bruk i offentlig sektor hvor også hensynet til definisjonsarbeid som prosess og ikke bare som et prosjekt ivaretas.

(2) Det anbefales å benytte URIs som identifikatorer på alle begreper. Det anbefales videre å utrede en beste praksis her basert på W3Cs anbefalinger om "Cool URLs for the Semantic Web".¹¹

(3a) Det anbefales å utvikle en egen helhetlig koordineringsprosess. Denne bør basere seg blant annet på ISO 860, men tilpasses til norsk offentlig sektor.

(3b) Det anbefales å utrede en distribuert utvikling og vedlikehold av begrepsapparater med deres livssyklus og innbyrdes sammenhenger.

6.1 *Anbefalinger relatert til anvendelsesområdene for definisjoner*

De følgende anbefalingene er relatert til en bestemt anvendelse av definisjoner. Vi tar her kun stilling til bruk av standardene og initiativene i kontekst av definisjonsarbeid, og eventuelle behov for ytterligere kartlegging innen disse anvendelsesområdene.

¹¹ <http://www.w3.org/TR/cooloris/>

Regelverk og rammebetingelser

(4) Det er viktig at det kommer fram hvorvidt en definisjon er ment uttrykt med nøyaktighet (for automatisering) eller unøyaktig for at anvendelsen skal være styrt av rettspraksis eller skjønn.

Det anbefales en spesiell utredning relatert til systemstøtte for beskrivelse av begreper og begrepsdefinisjoner under lovformulmuleringsprosessen. Dette bør omfatte beskrivelse av kontekst, referanser til rettskilder, tilsiktet nøyaktighet, gjenbruk mv.

Interne prosesser

(5) Begreper i intern forvaltning oppstår enten gjennom lov- eller forskifter eller gjennom fagdisipliner som ikke er styrt av lover og forskrifter (f.eks. generelt språk). Det er viktig her at det også dannes ordlister for fagdisiplinene. **Det bør vurderes om harmonisering av begreper på dette nivået bør helst gjøres i lys av en inkluderende metodikk med f.eks. wiki-verktøystøtte eller enn at sterke fagpersoner i miljøene får dominere.**

Meldingsutveksling

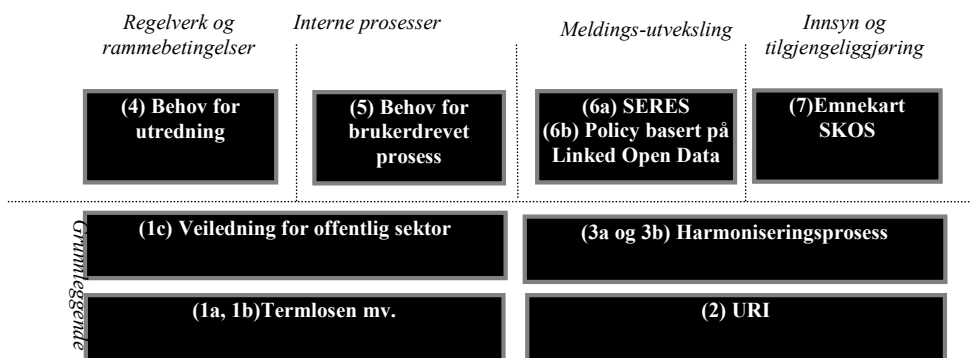
(6a) **Det anbefales å bygge videre på en infrastruktur rundt publisering av begreper og definisjoner basert på SERES som konsept.**

(6b) **Det anbefales å utrede en policy for publisering ikke bare av begreper, men åpne offentlige data (LOD) generelt på linje med hva som er blitt gjort i Storbritannia.** Denne må inneholde gode identifikasjons-anbefalinger for tilgjengeliggjøring av både begreper men også annet innhold.

Innsyn og tilgjengeliggjøring

(7a) **Det anbefales å bygge gjenfinningsløsninger på standardene ISO 13250 Topic Maps og W3C SKOS.**

(7b) **Det anbefales å ytterligere utrede policy for tilgjengeliggjørings av åpne data (se 6b)**



12

Figure 12 - Illustrasjon av anbefalinger

¹² Termlosen bygger på, og vi anbefaler i denne sammenheng ISO 704, ISO 1087, ISO10241 og ISO11179-4

BILAG A – Kriterier for definisjonsarbeid

BILAG B – Detaljert evaluering av standarder