



2004:22

Förstudierapport Standardmeddelande Bilaga 3:

Begreppsanalys och
informationsmodellering vid utveckling av
standardmeddelanden



Bilaga 3: Begreppsanalys och informationsmodellering vid utveckling av standardmeddelanden

Innehållsförteckning

1.	Sammanfattning	2
2.	Bakgrund	2
2.1	Varför modelleringsarbete?	3
2.2	Modellering vid utveckling av standardmeddelanden	3
3.	Begreppsanalys	5
3.1	Teoretisk grund	5
3.2	Praktiska exempel	6
3.2.1	Person	7
3.2.2	Personnummer	7
3.2.3	Namn	7
3.2.4	Adress	8
3.2.5	Datum	8
3.2.6	Inkomst	8
3.3	Begreppsdiagram och begreppsmodeller	8
3.4	Definitionsskrivande	10
3.5	Definiera eller inte definiera?	11
3.6	Tips och råd om begreppsanalys	11
4.	Informationsmodellering	12
4.1	Praktiskt exempel	12
4.2	Tips och råd om informationsmodellering	12
5.	Schemadesign	13
5.1	<i>Praktiska exempel</i>	<i>13</i>
5.1.1	Person	13
5.1.2	Personnummer	13
5.1.3	Namn	14
5.1.4	Adress	14
5.1.5	Datum	16
5.1.6	Inkomst	16
6.	Referenser och länkar	17

1.

Sammanfattning

Information och hantering av information är centralt för en organisations verksamhet. 24-timmarsmyndigheten och arbetet med en sammanhållen elektronisk förvaltning innebär samverkan och integration mellan myndigheter, bland annat genom utbyte av standardmeddelanden. Ett standardmeddelande är en definierad, ofta efterfrågad informationsmängd som utgör utdrag ur befintliga register, t.ex. en inkomstförfrågan. I ett standardmeddelande förekommer ett antal myndighetsgemensamma begrepp (t.ex. datum, personnummer). För att kommunikationen mellan uppgiftslämnare och mottagare inte ska leda till missförstånd är det viktigt att enas kring begreppens innebörd och relationerna mellan begreppen.

Denna rapport behandlar begreppsanalys, informationsmodellering och schemadesign. Dessa tre delar kan ses som steg i en arbetsprocess som går från verksamhetsbehov till teknisk realisering. Arbetsprocessen är nödvändig för att möjliggöra utbytet av standardmeddelanden och uppnå en effektiv sammanhållen elektronisk förvaltning.

Rapporten är beställd av Statskontoret och utformad av Metamatrix i samverkan med Terminologicentrum TNC.

2. Bakgrund

Statskontoret bjöd i juni 2004 in ett antal myndigheter till en workshop kring behovet av begrepps- och informationsmodellering vid utveckling av standardmeddelanden. Dagen innehöll föredrag och diskussioner där det framkom att det finns ett stort intresse för och behov av modelleringsarbete inom myndigheterna. Terminologi- och modelleringsarbete ses av myndigheterna som en förutsättning för att kunna börja interagera och elektroniskt utbyta information med varandra. Kompetensnivån inom detta område skiftar dock hos myndigheterna och stöd och främjandearbete efterfrågades av deltagarna.

Denna rapport syftar till att ge en introduktion till arbetsprocessen från existerande verksamhetsbehov till teknisk realisering. Rapporten berör några av de problem man kan stöta på vid definitionen av begrepp och vad man bör tänka på vid begrepps- och informationsmodellering. Dokumentet avslutas med ett kapitel om schemadesign, som illustreras med exempel.

Rapporten ska inte uppfattas som en handbok eller ett facit till modellering och schemadesign. De definitioner, modeller och XML-scheman som anges är exempel och inte färdiga lösningar. I verkligheten behövs betydligt mer

praktiskt arbete innan begreppen är färdigutredda för det sammanhang de ska användas i.

2.1 Varför modelleringsarbete?

Modelleringsarbete anses nödvändigt för det interna arbetet och är en förutsättning för att möjliggöra ett effektivt informationsutbyte mellan myndigheter. Deltagarna på workshopen var eniga om att särskilda insatser krävs för att tillgodose behovet av att enas kring innebörden av gemensamma begrepp (både internt och mellan myndigheter). Entydighet vad gäller termer och begrepp bidrar till ökad säkerhet och bättre kvalitet på informationen, vilket leder till att missförstånd kan undvikas.

Ett modelleringsarbete skapar härigenom förutsättningar för att förstärka servicen till medborgare genom ökad kundorientering, vilket är ett viktigt syfte med arbetet med 24-timmarsmyndigheten.

2.2 Modellering vid utveckling av standardmeddelanden

Arbetet med 24-timmarsmyndigheten medför krav på integration och samverkan mellan myndigheter. Ett exempel på informationsutbyte mellan myndigheter är de så kallade standardmeddelandena. Med standardmeddelande avses en definierad informationsmängd som utgör utdrag ur existerande register och som frekvent efterfrågas av ”många” vid handläggning av olika ärendetyper, eller indata som frekvent skall lämnas av ”många” till existerande register och som:

är formaterat och strukturerat enligt fastlagda normer

är tillgängligt enligt fastlagda nyttjanderegler

finns publicerat (inklusive nyttjanderegler) i en öppen och allmänt tillgänglig katalog på webben

Syftet med standardmeddelanden är att göra det så lätt som möjligt för förvaltningens organisationer att få tillgång till sådan information som frekvent efterfrågas i samband med handläggningen av ett enskilt ärende. Det ska vidare vara enkelt att tillgodose behovet av sådan information. På motsvarande sätt ska det vara möjligt för uppgiftslämnande organisationer att utveckla effektiva lösningar för att fullgöra sin uppgiftsskyldighet och

för mottagande myndigheter att utveckla enhetliga lösningar för att ta emot uppgifter från flera olika uppgiftslämnare.¹

Vid utbytet av standardmeddelanden måste rätt information hämtas ur rätt register och bearbetas på ett väldefinierat sätt. Annorlunda uttryckt: begreppen bakom orden eller termerna måste vara väldefinierade och dessutom måste regler finnas för hur den tillhörande informationen ska anges (termer, siffror etc.). När "datum" ska anges måste rätt tidsuppgift hämtas in och sedan anges på ett standardiserat sätt. En analys av informationen på begreppsnivå är en nödvändig förutsättning för senare regelskrivning av information på datanivå.

Det finns ett stort antal myndighetsgemensamma begrepp som kan förekomma i standardmeddelanden, under workshopen uppskattades denna mängd till cirka 30-100 st. Exempel på sådana begrepp är: Ärende, dokument, handling, beslut, datum, akt, journal, adress, person, personnummer, företag, myndighet, namn, inkommet, medborgare, patient, arkiv, sekretess, blankett, möte, inkomst.

I den här rapporten kommer sex av dessa begrepp behandlas närmare för att exemplifiera resonemangen kring begreppsanalys, modellering och schema-konstruktion.

person
personnummer
namn
adress
datum
inkomst

Dessa sex får illustrera möjliga begrepp i ett standardmeddelande och kommer att användas som exempel för att belysa modelleringsprocessen. Begreppen kan ses som allmänna och vanligt förekommande i framtida standardmeddelanden som ska utväxlas mellan myndigheter. Vid ett standardmeddelande för inkomstförfrågan så är exempelvis *personnummer*, *datum* och *inkomst* intressanta.

Man kan identifiera tre steg i arbetsprocessen modellering vid utveckling av standardmeddelanden:

Begreppsanalys
Informationsmodellering
Schemadesign

¹ <http://www.24-timmarsmyndigheten.se/>

Dessa steg kommer att behandlas närmare nedan, med fokus på begreppsanalys (inklusive konstruktion av begreppsmodell eller begreppsdiagram och definitionsskrivande).

3. Begreppsanalys

Behovet av begreppsanalys uppstår när man upptäcker att kommunikationen i tal och skrift är oklar och behöver förtydligas. Tvetydiga termer och luddiga definitioner skapar problem när instruktioner lämnas, när avtal skrivs eller när data ska utbytas mellan system och sedan kanske också tolkas av människor. I många kommunikationssituationer – inte minst vid utbyte av standardmeddelanden mellan myndigheter – krävs att data mer eller mindre automatiskt hämtas och redovisas ur register. Man måste noga understryka vikten av att dessa data är strukturerade och lagrade på ett standardiserat vis och att de begrepp som data beskriver är väldefinierade. Dessutom måste regler finnas för på vilket sätt data rent konkret ska registreras.

En väl genomförd begreppsanalys är en nödvändig förutsättning för strukturering av data. I begreppsanalysen ringar man in centrala begrepp och utreder vad något är och vad det inte är. Man drar också upp gränserna mellan olika begrepp – hur de förhåller sig till varandra och vad som är skillnaden mellan näraliggande begrepp. Begreppens inbördes relationer kan illustreras grafiskt tillsammans med de termer som representerar begreppen i begreppsdiagram eller begreppsmodeller, och de kan också beskrivas språkligt i definitioner.

I många sammanhang vill man också analysera vilka begrepp man behöver hantera information om och vilken denna information är. Man brukar då arbeta fram en så kallad informationsmodell (se vidare avsnitt 4), som kan bygga på en begreppsmodell, men som även kan utvecklas parallellt med begreppsanalysen.

3.1 Teoretisk grund

Här ges inte någon djup genomgång av terminologins teori, vilken ligger till grund för begreppsanalysen, men för att kunna följa resonemanget bör man ändå ha skillnaden mellan *term* och *begrepp* klar för sig. *Term* och *begrepp* är nämligen inte synonyma i detta sammanhang, vilket de ofta är i allmänt språkbruk.

Termen är det språkliga uttrycket och *begreppet* är det som uttrycket står för.

Det är också nödvändigt att känna till de tre huvudtyperna av begreppsrelationer som man arbetar med i begreppsanalys, nämligen *generiska*, *partitiva* och *associativa* relationer.

Generiska relationer är "typ av"-relationer (*lön* är en typ av *lön*).
Partitiva relationer är "del av"-relationer (en *stam* är en del av ett *träd*).
Associativa relationer är de relationer som varken är generiska eller partitiva (*träd* växer i en *skog*, av *träd* kan man tillverka *papper* etc.)

3.2 Praktiska exempel

Betraktar man de sex termerna *person*, *personnummer*, *namn*, *adress*, *datum* och *inkomst* ser man tydligt att relationerna mellan begreppen varken är generiska eller partitiva relationer (*namn* är inte en "typ av" *person*; *inkomst* är inte en "del av" *datum*, etc.). I stället hör de ihop på ett associativt sätt, dvs. i ett visst sammanhang behöver de hanteras i förhållande till varandra. Man kan till exempel tänka sig att en person *har* alla de övriga begreppen utom *datum*. Men eftersom den viktigaste uppgiften med begreppsanalysen är att ringa in vad begreppen är, räcker det inte att beskriva de associativa förhållandena till de andra begreppen. Man måste i stället titta närmare på varje begrepps kärna. Det är då lämpligt att för varje begrepp ställa sig följande frågor:

Vad är det för egenskaper som gör att en viss företeelse eller ett visst föremål omfattas av begreppet (vad gör en stol till en stol)?

Finns det andra begrepp som det är viktigt att avgränsa begreppet mot?

Finns begreppet redan definierat i relevanta källor?

Finns det någon term för begreppet, eller kan det endast beskrivas med en definition eller förklaring?

När det finns en term i bruk för begreppet, finns även följande frågor att ställa sig:

Används termen i en så övergripande betydelse att den blir oprecis i sammanhanget?

Är den valda termen för precis, så att den utesluter sådant man vill att den ska omfatta?

Är termen olämplig av något annat skäl, t.ex. på grund av att den avviker från etablerat språkbruk?

Olika begrepp innebär olika svårigheter, vilket, åtminstone i detta exempel, bland annat beror på vilken grad av precision de används med i allmänspråket. Det gör också att frågorna ovan får olika tyngd för olika begrepp. Nedan presenteras en rimlig inledande begreppsanalys för vart och ett av de aktuella begreppen. Observera att begreppsanalysen inte är fullständig för något av begreppen, främst tas relevanta frågeställningar för varje begrepp upp. I några fall nämns även funderingar och överväganden som kan bli aktuella i samband med en informationsmodell snarare än i samband med en begreppsmodell eller ett begreppsdiagram. Det är inte säkert att samma begrepp behöver ingå såväl i begreppsdiagram eller begreppsmodell som i motsvarande informationsmodell.

3.2.1 Person

Menar man *fysisk person* eller *juridisk person* eller både och? Om man menar *fysisk person* – framgår det då så tydligt av sammanhanget att termen inte behöver ändras till *fysisk person*, eller finns det skäl att ändå göra det för säkerhets skull? Möjligen räcker det att ange att ”med person avses fysisk person”. Detta behöver då inte anges i en ordlista utan kan i stället preciseras i anvisningar eller regelverk. I en informationsmodell skulle *person* troligen beskrivas med en rad attribut (personnummer, namn etc.), som i begreppsmodellen är separata begrepp med egna definitioner.

3.2.2 Personnummer

Personnummer har definierats som ’enhetligt utformat person-id registrerat i folkbokföringen’². Av definitionen kan man ana att det finns flera olika *person-id*. I samma källa kan man läsa att *person-id* är en ’identitetsbeteckning för fysisk person’ och att det förutom *personnummer* bl.a. finns *reservnummer* och *samordningsnummer*. Det är viktigt att analysera om det egentligen är *person-id* man är ute efter snarare än *personnummer* – annars riskerar man problem med att hantera information om personer som inte har svenska personnummer? Om man har bestämt sig för att hantera även juridiska personer, måste man även ta med begreppet *organisationsnummer* i begreppsanalysen.

3.2.3 Namn

Vid termen *namn* måste man fråga sig vilken typ av namn man är ute efter: *personnamn*, *förnamn*, *efternamn*, *mellannamn*, *tilltalsnamn*, *produktnamn*, *organisationsnamn* osv. Det är viktigt att undersöka hur dessa begrepp definieras och hur de förhåller sig till varandra. Används ens dessa termer?

² Enl. Socialstyrelsens InfoVU-projekt, Basinfo, se <http://www.sos.se/Hs/storproj/inf.htm>

Förekommer andra? I en informationsmodell kan det hända att *namn* behöver delas upp i flera underkategorier.

3.2.4 Adress

Precis som för *namn* måste man bestämma sig för vilken *adress* som avses: *utdelningsadress*, *postortsadress*, *postadress*, *postal adress*, *utlämningsställe*, *folkbokföringsadress*, *e-postadress* eller något annat? Vad står det enkla *adress* egentligen för? Vad står de övriga termerna för? Används de i dag, och hur förhåller sig begreppen till varandra? Vilka standarder gäller?

3.2.5 Datum

Begreppet *datum* i sig ställer i allmänhet inte till några kommunikationsproblem – det är en angivelse av en viss dag ett visst år. Det viktiga är i stället även här att fastslå vilket *datum* som ska anges i ett visst sammanhang. Är det t.ex. det datum då standardmeddelandet skickades, det datum då förfrågan inkom, det datum då ärendet registrerades? Begreppet behöver troligen inte definieras, bara anges med ett tydligare uttryck. Eventuellt behöver även närliggande begrepp såsom tidsperiod (från och med datum och till och med datum) läggas till. Någonstans i dokumentationen – men inte i ordlistan – bör också anges i vilken form datum ska skrivas (2004-07-06 eller 6 juli 2004 eller 6/7-04 etc.). Lämpligen följs SS-ISO 8601:s rekommendationer om skrivsätt för datum (20040706 eller 2004-07-06).

3.2.6 Inkomst

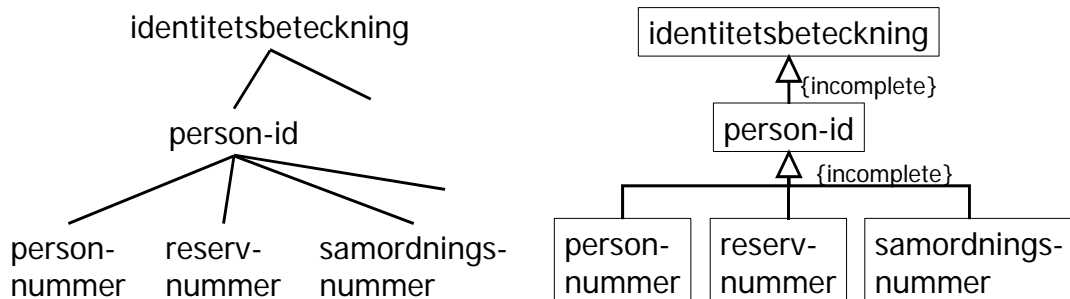
Termen *inkomst* är troligen den av exempeltermerna som kan ställa till mest problem även i kommunikation mellan människor. Vad är egentligen *inkomst*? Vilka typer av ekonomisk ersättning ingår i *inkomst*? Hur skiljer sig begreppet från *lön* och *intäkt*? Avses *taxerad inkomst* eller någon annan typ av *inkomst*?

3.3 Begreppsdiagram och begreppsmodeller

För att åskådliggöra begreppens relationer till varandra kan man grafiskt rita in dem i ett begreppsdiagram eller en begreppsmodell. Den grafiska presentationen skapar överblick över begreppen och deras relationer, och är därför ofta ett bra stöd i begreppsanalysen.

Ofta används begreppsdiagram som presentationsform (se ISO 1087 och ISO 704). Det finns dock ingenting som hindrar att man i stället använder ett modelleringsspråk, t.ex. UML, för själva notationen. Man kan låta den tänkta målgruppen avgöra vilket notationssätt man väljer. Till exempel kan det vara lämpligt att välja ett modelleringsspråk om man ska fortsätta arbetet med att konstruera en informationsmodell eller en datamodell, eftersom man då kan fortsätta arbetet med samma notation (t.ex. UML). Är målet en

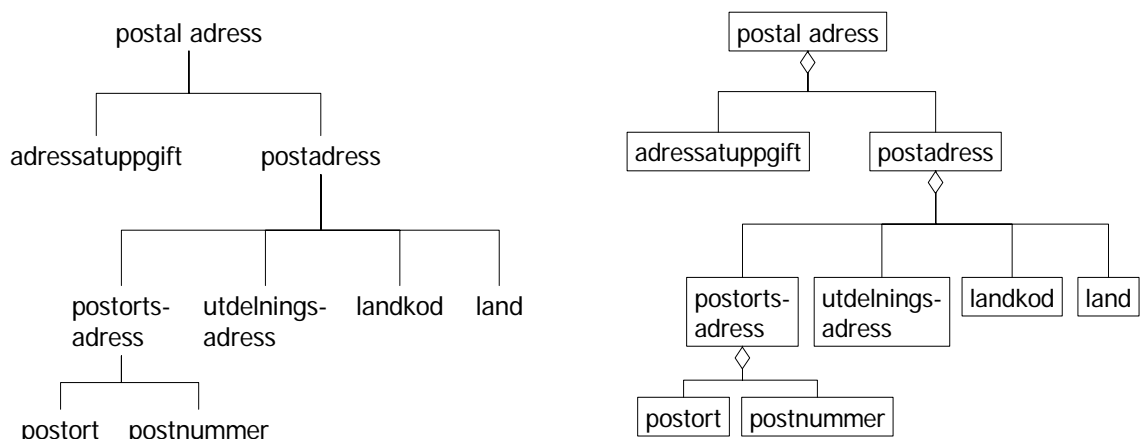
ordlista kan det vara enklare med begreppsdiagram. I exemplen i figurerna är både begreppsdiagrammen och begreppsmodellerna en aning förenklade. Diagrammen och modellerna bör heller inte anses vara färdiga – vidare begreppsanalys krävs för att förbättra dem.



Figur 1: Relationerna mellan begrepp kan presenteras grafiskt med ett begreppsdiagram (till vänster) eller en begreppsmodell i t.ex. UML (till höger). I exemplet visas generiska relationer.

De sneda strecken i begreppsdiagrammet och de uppåtriktade pilspetsarna i begreppsmodellen i figur 1 symboliserar generiska relationer (*person-id* är en typ av *identitetsbeteckning*). De ”tomma” strecken i begreppsdiagrammet och {incomplete} i begreppsmodellen symboliserar att det kan finnas fler typer av identitetsbeteckningar respektive person-id.

Man kan också symbolisera partitiva relationer och associativa relationer med båda notationssätten.



Figur 2: Partitiva relationer presenterade som begreppsdiagram (till vänster) och begreppsmodell (till höger).

I figur 2 symboliserar de raka strecken i begreppsdiagrammet och parallelogrammen i begreppsmodellen ”del av”-relationer (notera att exemplet med de olika adresserna är hämtat ur SS-61 34 01 från 1995, som nu är anpassad till en europeisk standard och ersatt av SS-1 34 01 från



2003).

Figur 3: Associativ relation presenterad som begreppsdiagram (till vänster) och begreppsmodell (till höger).

I figur 3 symboliserar pilen relationen. Ofta skriver man också ut vilken relation det handlar om. I begreppsmodellen skriver man också oftast ut multipliciteten, dvs. antalsförhållandet: en person har ”noll till många” person-id, dvs. en person kan sakna person-id, men kan också ha flera.

3.4 Definitionsskrivande

Resultatet av begreppsanalysen bör också omfatta textuella definitioner av begreppen. En definition är en språklig beskrivning av begreppet, exempelvis:

dörr - öppningsbar anordning för tillslutning av en öppning i en vertikal avskiljande konstruktion mellan beträdbara utrymmen, avsedd för genomgång av gående personer i mer eller mindre upprätt ställning

Definitionen ger här en bild av begreppet *dörr* genom att räkna upp en unik kombination av kännetecken, dvs. ”det som gör en dörr till en dörr”. Definitionen avgränsar också begreppet *dörr* mot andra näraliggande begrepp, som *lucka*, *port*, och *grind*.

En bra definition är varken för vid eller för snäv, dvs. den innefattar inte sådant som inte omfattas av begreppet, men den utesluter inte heller sådant som omfattas av begreppet. Definitionen av begreppet *personnummer* får därför inte gälla även för *reservnummer*. En definition av det överordnade begreppet *person-id* måste däremot omfatta både *personnummer* och *reservnummer*.

Man kan le åt definitioner som på ett tillkrånglat sätt förklarar något enkelt och välkänt, som *dörr*. Men, som tidigare sagts, ibland behövs en definition inte bara för att beskriva vad något är utan också för att man ska kunna avgöra vilka företeelser som inte omfattas av begreppet (jämför också t.ex. *namn* eller *adress* ovan).

Ibland ifrågasätts nyttan med definitioner, och begreppsmodeller eller begreppsdiagram påstås ge en tydligare bild av begreppen. Det är sant att grafiska presentationer kan ge en bra överblick, men definitioner är ofta en

bättre och nödvändig källa för förståelse av begreppen. En definition av ett begrepp kan användas självständigt i olika sammanhang, även när hela systemet inte presenteras.

I samband med utvecklingen av standardmeddelanden syftar definitionerna till att ge snabbare och klarare upplysningar om vad som avses med varje term. En förutsättning för det elektroniska utbytet är att informationen tolkas på samma sätt av respektive part.

3.5 Definiera eller inte definiera?

I begreppsanalysen ingår också att analysera vilka begrepp som behöver definieras och vilka som snarare behöver någon annan typ av precisering. Termer och definitioner brukar samlas i en ordlista som kan vara ett självständigt dokument eller ingå som en del i ett annat dokument. Uppgifter som inte beskriver själva begreppet, t.ex. anvisningar om hur datum ska skrivas eller hur personuppgifter ska hanteras, hör inte hemma i definitionen. I stället kan sådana anvisningar samlas i ett separat dokument.

3.6 Tips och råd om begreppsanalys

När man arbetar med begreppsanalys bör man ha följande i tankarna:
Arbeta inte ensam

Se till att arbetsgruppen består både av personer som är väl insatta i fackområdet eller verksamheten och av terminologer

Tänk igenom vilka begrepp som verkligen behöver definieras och vilka som snarare behöver preciseras på annat sätt

Leta efter redan gjorda definitioner ur relevanta källor

Ta reda på om termen används bland språkbrukarna inom fackområdet – och om den används i den avsedda betydelsen

4. Informationsmodellering

Syftet med informationsmodellering är att strukturera den information som en verksamhet måste kunna hantera om olika företeelser. Informationsmodellering i samband med systemanalys fokuserar på den informationsmängd som det specifika systemet ska arbeta med. Detta innebär oftast att endast en delmängd av en verksamhets kompletta begreppsmodell ingår i informationsmodellen för systemet. Dessutom kan man under informationsmodelleringen komma att tillföra nya begrepp. Informationsmodelleringen behöver alltså inte göras efter begreppsmodelleringen utan de båda kan utvecklas parallellt.

4.1 Praktiskt exempel

Om man exempelvis ska modellera ett standardmeddelande för inkomstförfrågan så är troligen inte begreppen namn eller adress så relevanta. Däremot är det intressant att titta närmare på och detaljera begreppen inkomst, datum och personnummer.

Inkomst bör ha delats in i olika typer av inkomst (inkomst av tjänst, inkomst av kapital osv.) redan i den övergripande begreppsmodellen. För datum kan man behöva definiera olika typer av datum. Eventuellt behöver även närliggande begrepp såsom tidsperiod (från och med datum och till och med datum) läggas till. För personnummer måste man besluta om huruvida det ska användas som person-id. Här kan man exempelvis komma att avvika från verksamhetens definition då man för systemet måste ta hänsyn även till omgivande systems krav inklusive säkerhetskrav. Verksamheten kanske använder personnummer som person-id (i begreppsmodellen) men systemet förväntas av integritetsskäl använda ett mer anonymiserat person-id vid kommunikationen med omvärlden.

4.2 Tips och råd om informationsmodellering

I en organisation som har en dokumenterad begreppsmodell är det lämpligt att relatera informationsmodellen till begreppsmodellen och vid behov komplettera begreppsmodellen och göra den mer detaljerad. Här bör man använda sunt förnuft vid valet av vilka termer som tas med i begreppsmodellen. Om man vid systemanalysen exempelvis har definierat ett antal datumbegrepp för att underlätta intern spårning och loggning så kanske inte dessa behöver inkluderas i verksamhetens begreppsmodell. Det kan räcka med att dessa finns dokumenterade i informationsmodellen. Begreppsmodellen syftar till att underlätta kommunikation mellan olika delar i

verksamheten, den ska ses som en levande modell med information och begrepp som ändras och tillkommer. Med tanke på detta kan det vara lämpligt att hålla begreppsmodellen relativt okomplicerad, för att underlätta kompletteringar och ändringar.

5. Schemadesign

När man beskrivit informationsmodellen (ofta grafiskt med t.ex. notations-språket UML) är det dags att börja se hur denna ska implementeras. Informationsmodellen implementeras i alla delar av ett system; i programkod exempelvis i form av klasser, i databaser exempelvis i form av tabeller och kolumner och i XML-scheman i form av element och attribut.

En generell regel avseende schemadesign är att först utreda om andra redan konstruerat ett schema för det aktuella begreppet. I så fall är det i princip bäst att återanvända något av dessa.

Vid översättning av en informationsmodell till konkreta XML-scheman är arbetet med att identifiera gemensamma datatyper och element den mest tidskrävande aktiviteten.

5.1 Praktiska exempel

I nedanstående exempel illustreras några av de designbeslut man kan ställas inför under schemadesignen. Exempelen utgår från det tänkta standard-meddelandet inkomstförfrågan.

5.1.1 Person

Som begrepp representerar person den för vilken en inkomstförfrågan ska göras. I schemat räcker det med att låta personen representeras av ett unikt person-id. Däremot är det troligt att begreppet person får en mer komplett beskrivning i den databas som ska hålla ordning på informationen om personen.

5.1.2 Personnummer

Personnummer kan i systemet representera en persons personnummer men kan också vara den unika identifieraren (person-id) för just den personen. Oavsett vilken roll personnumret ges är det lämpligt att schema-beskrivningen inkluderar en mer detaljerad syntaxkontroll.

Nedan exemplifieras en variant på implementation med hjälp av begränsningsmönster för XML-schemats inbyggda datatyp string.

```
<xs:simpleType>
  <xs:restriction base="xs:string">
    <xs:pattern value="(19|20)[0-9]{6}-[0-9]{4}"/>
  </xs:restriction>
</xs:simpleType>
```

Det tillåtna mönstret anger att personnumret måste börja på "19" eller "20", som följs av sex siffror mellan 0-9, ett streck och avslutas med fyra siffror mellan 0-9. Genom att i schemat begränsa mönstret säkerställer man att personnumret har korrekt format.

5.1.3 Namn

Detta begrepp är inte aktuellt i det tänkta standardmeddelandet inkomstförfrågan men exempel på strukturer för namnelement hittas exempelvis hos hr-xml (www.hr-xml.org), som även tillhandahåller dokumentation. Begreppet namn är ett sådant som redan blivit designat av många. Det kan därför vara bra att inleda med att göra en inventering av gjorda definitioner och strukturer.

5.1.4 Adress

Vilka olika adresser (postadress, besöksadress, leveransadress etc.) som ska stödjas bör ha framkommit under informationsmodelleringen. Centrala frågeställningar som kan kvarstå vid schemakonstruktionen är bland annat hur nedbruten en adress ska vara i sina beståndsdelar. Till exempel om man hanterar en leveransadress endast i syfte att skriva ut den i fält på en blankett så kan det räcka med en fritextsträng som datatyp i schemat, men om leveransadressen ska användas av ett system som ska räkna fram geokoordinater för en specifik adress så behöver kanske till och med utdelningsadressen separeras i delar som gata och gatunummer.

Vidare behöver man bestämma sig för om alla olika adresselement ska implementeras med en gemensam datatyp eller om man ska ha flera olika datatyper för adresser. Om man har en detaljerad adressdatatyp kan det även vara intressant att skapa en vokabulär för postorter. Vokabulärer kan beskrivas på flera olika sätt. Ett sätt är att förse elementet med ett attribut i vilket användaren kan ange vilken vokabulär man använt för innehållet i elementet.

```
<xs:complexType name="PostOrtTyp">
  <xs:simpleContent>
    <xs:extension base="xsd:string">
      <xs:attribute name="vokabulaerNamn"
        type="xsd:string"/>
```

```

        use="optional"/>
    </xs:extension>
</xs:simpleContent>
</xs:complexType>

```

Om man vill kräva angivande av vokabulärnamn kan man använda use="required" istället för use="optional".

För kortare vokabulärer kan man ange en lista av acceptabla värden att använda. Om man exempelvis enbart skall adressera kommuner i Stockholms län så kan man skapa följande postortselement.

```

<xs:element name="postOrtIStockholmsLan">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:enumeration value="Botkyrka"/>
      <xs:enumeration value="Danderyd"/>
      <xs:enumeration value="Ekerö"/>
      <xs:enumeration value="Haninge"/>
      <xs:enumeration value="Huddinge"/>
      <xs:enumeration value="Järfälla"/>
      <xs:enumeration value="Lidingö"/>
      <xs:enumeration value="Nacka"/>
      <xs:enumeration value="Norrtälje"/>
      <xs:enumeration value="Nykvarn"/>
      <xs:enumeration value="Nynäshamn"/>
      <xs:enumeration value="Salem"/>
      <xs:enumeration value="Sigtuna"/>
      <xs:enumeration value="Sollentuna"/>
      <xs:enumeration value="Solna"/>
      <xs:enumeration value="Stockholm"/>
      <xs:enumeration value="Sundbyberg"/>
      <xs:enumeration value="Södertälje"/>
      <xs:enumeration value="Tyresö"/>
      <xs:enumeration value="Täby"/>
      <xs:enumeration value="Upplands Väsby"/>
      <xs:enumeration value="Upplands-Bro"/>
      <xs:enumeration value="Vallentuna"/>
      <xs:enumeration value="Vaxholm"/>
      <xs:enumeration value="Värmdö"/>
      <xs:enumeration value="Österåker"/>
    </xs:restriction>
  </xs:simpleType>
</xs:element>

```

5.1.5 Datum

I XML-schemastandarden finns ett antal datum och tidsdatatyper som baserar sig på ISO 8601. Vid schemadesign behöver man besluta sig för vilken av dessa som ska användas, alternativt om man kommer att använda en fritextsträng för angivande av datum.

5.1.6 Inkomst

Givet att man vid informationsmodelleringen har bestämt sig för om eventuell nedbrytning i olika inkomsttyper ska göras så bör schemakonstruktionen vara ganska enkel, då en inkomst rimligen representeras med ett antal siffror. Detta såvida man inte behöver kunna ange inkomst i olika valutor. Om så är fallet kan samma typer av vokabulärsdefinitioner som beskrevs för adresser även bli aktuella här. Antingen anger man vilken vokabulär man använt för elementet eller så definierar man en lista av acceptabla värden att använda. I detta fall skulle listan utgöras av godtagbara valutor istället för postorter.

6. Referenser och länkar

För mer information om begreppsanalys och definitionsskrivande, se: ISO 704: Terminology work – Principles and methods ISO 15188: Project management guidelines for terminology standardization ISO 1087-1: Terminology work – Vocabulary – Part 1: Theory and application SS-ISO 10241: Internationell terminologistandard – Utarbetande och utformning

Samtliga standarder kan beställas hos SIS, Swedish Standards Institute (www.sis.se).

Terminologiguiden – en introduktion till terminologiarbete i teori och praktik (2004): En kortfattad guide till terminologiarbete, baserad på ISO-standarderna 704, 860 och 10241. Översatt och bearbetad från det engelskspråkiga originalet Guide to terminology (2 utg., 2001). Beställs hos Terminologicentrum TNC (www.tnc.se)

Standarder och handböcker om hur UML kan anpassas för och användas i begreppsanalys har vi inte identifierat, dock pågår arbete inom området.

För mer information om informationsmodellering se: www.sis.se – anordnar bl.a. kurser i hur UML kan användas vid informationsmodellering

För mer information om XML-scheman och schemadesign se: www.w3.org/XML/Schema - World Wide Web Consortium, om XML-schema som Recommendation www.xml.com (Se t.ex. <http://www.xml.com/pub/a/2004/04/28/ubl.html> om UBL, Universal Business Language) <http://www.xfront.com/BestPracticesHomepage.html> - Schemadesign, dokumentation, tutorials http://xml.gov/documents/in_progress/developersguide.pdf - XML Developer's Guide (för schemadesign se kapitel 4)