

Subproject: *Supporting facilities*

*Tema Informationssamverkan:*

# **Begreppsmodelldriven informationssamverkan – utbytesansats <sup>1</sup>**

**Version 0.8  
(2004-05-05)**

**Stig Berild**

**(Santa Anna IT Research Institute AB)**

---

<sup>1</sup> Rapport framtagen i samarbete mellan e-Society II-projektet ([www.skriver.nu/esociety](http://www.skriver.nu/esociety)) vid Santa Anna IT Research Institute ([www.santaanna.se](http://www.santaanna.se)) och Serviam-projektet ([www.serviam.se](http://www.serviam.se)) samt med stöd från Stiftelsen SISU och Vinnova.

## 1 Inledning

Rapporten ingår i en rapportserie under temat Informationssamverkan (semantik, informationsmodellering och informationsutbyte) med relatering till SOA (Service Oriented Architectures). För en aktuell introduktion till såväl denna rapport som rapportseriens övriga rapporter hänvisas till ”Tema Informationssamverkan – Rapportserieöversikt” ([www.skriver.nu/esociety](http://www.skriver.nu/esociety) eller [www.serviam.se](http://www.serviam.se)).

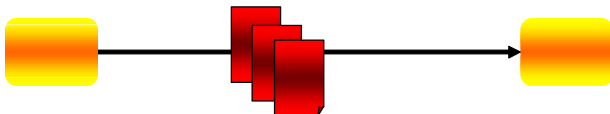
Se rapporten endast som ett diskussionsinlägg och några funderingar kring informationssamverkan relaterat till informationsutbyte. Säkerligen innehåller texten ett antal ”grodor” och oövertänkta funderingar. Läs med andra ord rapporten med denna omständighet för ögonen.

## 2 Informationssamverkan – förutsättningar och principer

Detta kapitel diskuterar de möjligheter och problem kring begreppsmodellbaserad samverkan som följer av olika förutsättningar. Diskussionen utgår i första hand från generella samverkansbehov mellan två intressenter för att inte onödigtvis komplicera resonemangen. Fler inblandade intressenter innebär ofta fler hänsynstaganden och kanske besvärligare problem även om den grundläggande problematiken kvarstår.

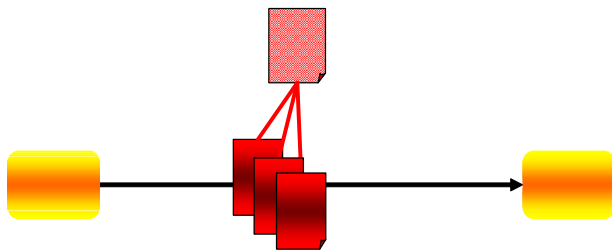
### 2.1 Konventionellt XML-utbyte

Först några ord om normalt utbyte med hjälp av XML. I sin enklaste form utbyts XML-dokument, dvs. informationen inklusive de märkord som är satta att förklara innehållet.



Figur 1

Affärsmässigt och annat utbyte under ordnade former kräver normalt en överenskommelse om vilka typer av dokument som ska utbytas. För varje typ av dokument specificeras vilka typer av informationselement som får förekomma och under vilka villkor. Dessutom specificeras informationselementens ingående ordning i form av en hierarkisk struktur. En dokumenttypspecifikation kallas i standarden XML 1.0 för en Document Type Definition (DTD). Om istället standarden XML Schema används för samma ändamål kallas typspecifikationen för *schema* rätt och slätt. Båda standarderna har i detta avseende samma syfte. Skillnaden är att XML Schemastandarden har betydligt rikare uttrycksmöjligheter. En introduktion till XML Schema ges i ”XML Schema – en översikt”.



Figur 2

Varje XML-dokument har därefter möjlighet att peka på det schema som gäller för dokumentet ifråga. Eller snarare, dokumentet påstår att det är utformat i enlighet med reglerna för schemat. Det är upp till mottagaren att kontrollera att så är fallet i realiteten.

Självfallet kan två parter komma överens om ett antal scheman som acceptabla att använda vid olika typer av informationsutbyten. För varje part kan det i det generella fallet bli ganska mycket att hålla reda på om utbyte ska kunna äga rum mellan många parter och för många syften. De scheman som gäller för ett visst avgränsat syfte eller tillämpningsområde brukar kallas vokabulär. Med hjälp av Namespaces-standarderna kan olika informationselement i olika scheman referera till varandra för att indikera överensstämmelser mm. Dock ställs inga sådana samordningskrav. Inte heller är det säkert att referenserna pekar exakt rätt eller ens på något som alls kan ge vägledning.

## 2.2 Gemensam begreppsmodell

Dags att föra in begreppsmodeller ”på arenan”.

Två intressenter A och B önskar samverka. En första förutsättning är att intressenterna kommer överens om vad som ska informationsutväxlas. A och B måste komma överens om en gemensam syn på den verklighet man önskar utbyta information om utgående från överenskommen avgränsning av verkligheten och med hänsyn till överenskommet syfte. De behöver helt enkelt definiera en gemensamt accepterad begreppsmodell. Denna begreppsmodell innehåller de begrepp som får användas samt vad respektive begrepp exakt betyder. Därtill innehåller modellen information om de begreppsband som anses angelägna att kunna hantera för att beskrivningen av verkligheten ska anses godtagbart representera den syftesavgränsade verkligheten.

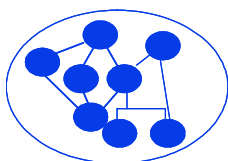
Begreppsmodellens omfattning beror alltså på syfte, samsyn, hanterbarhet, med mera.

Begreppsmodellen utgör en - förhoppningsvis för A och B gemensamt välavpassad - spegelbild av denna verklighet. Som sådan reglerar begreppsmodellen vilka informationselement som kan utbytas liksom ett antal villkor som är förknippade med innehållet och dess struktur.

Nåväl, först måste A och B sätta sig vid förhandlingsbordet för att komma överens om begreppsmodellens innehåll. Redan där kan det bli tvärstopp, om det vill sig illa. Svenska språket, liksom alla andra generella språk ger utrymme för tvetydigheter, olika tolkningar.

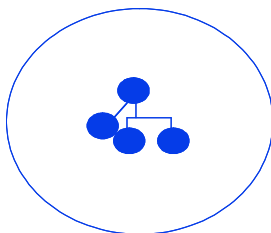
Här finns en uppenbar risk för att intressenterna pratar förbi varandra eller missförstår. Ett mer begränsat språk anpassat för att just dokumentera begreppsmodeller med önskad precision behövs. Jämför två elektriker som tillsammans ska utföra ett installationsarbete. De är utbildade på att använda och förstå det standardiserade språk som används för att dokumentera elscheman. Använder de elschemat som referenspunkt under diskussionen om hur installationen bäst ska lösas förstår de varandra exakt och direkt. Skulle de istället prata med varandra på vanlig svenska skulle risken för missförstånd bli uppenbar. Eller, ta två musiker som önskar spela duo men som spelar efter två olika partiturnotationer. Det låter nog inte så vackert (om inte någon med kunskap om båda sett till att harmonisera dem förstås).

I det aktuella fallet behöver bara A och B komma överens vilket ökar sannolikheten för en intrimmat effektiv modell.



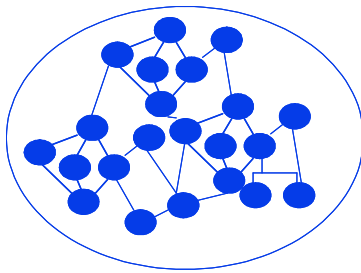
Figur 1

I det generella fallet, när antalet intressenter sannolikt är många fler, kanske med en global spridning och olika kulturella och affärsmässiga perspektiv på samverkan, kan det snabbt bli betydligt mer komplicerat med stridande viljor, urvattnade kompromisser, mm. I det läget kan det vara en framkomlig väg att inledningsvis endast försöka komma överens om det centrala och därefter allteftersom komplettera på basis av erfarenhet och/eller trängande behov.



Figur 2

I andra fall är det precis tvärtom. En branschorganisation eller standardiseringsorgan har under möda och över lång tid lyckats av väl täcka in "sitt" tillämpningsområde, resulterande i en omfattande begreppsmodell inom vilken det sedan är fritt fram för olika konstellationer av intressenter att välja och vraka för sina speciella ändamål. Ibland fungerar detta väl, till exempel inom väl avgränsade etablerade branscher där samverkan alltid pågått men tidigare i andra former. Ett annat exempel är samverkansparter som känner varandra rimligt väl och utgör en förhållandevis stabil gruppering och där behoven varit trängande – ekonomi, prestanda, konkurrenskraft, mm driver på. I andra fall blir denna begreppsmodell ett monument att kanske främst hämta inspiration och kunskap ifrån men vars innehåll inte hänger med i utvecklingen eller tagits fram utan tillräcklig lyhördhet för intressenternas behov.



Figur 3

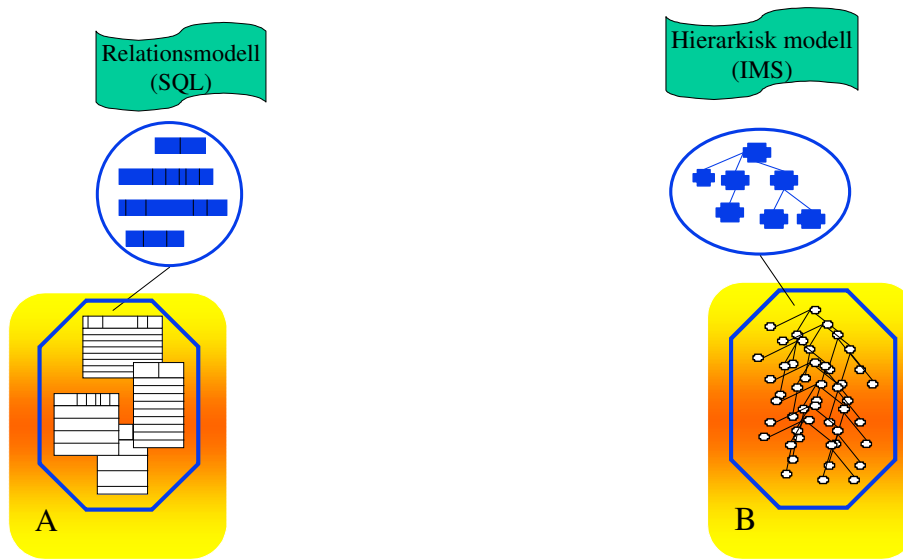
Denna problematik är ingalunda unik i samband med begreppsmodeller. Precis samma problemsituationer uppstår när olika grupperingar av intressenter har att komma överens om en lämplig uppsättning meddelandetyper (vokabulär) att användas vid informationsutbyte.

För begreppsmodeller finns ett otal språk att tillgå, allt från en enkel hierarkisk notation över binära språk till Unified Modeling Language (UML) eller något än mer avancerat (till exempel för att uttrycka så kallade ontologier inom Artificiell Intelligens). Vilket språk som är att föredra beror på aktuellt behov av uttryckskraft, vad de som har att arbeta med modellen är vana att använda, vilka krav som eventuella teknikstöd ställer, med mera. Att säga att ett visst språk "är bäst" är med andra ord ganska meningslöst. Bedömningen måste alltid göras utifrån ett givet sammanhang och dess specifika förutsättningar.

Med hjälp av ett modelleringsspråk står det nyttjaren fritt att formulera vilka begreppsmodeller som helst som går att uttrycka med det aktuella språket. En och samma partiturnotation kan ju användas för att skapa vilka musikverk som helst.

### **2.3 Olika interna begreppsmodeller**

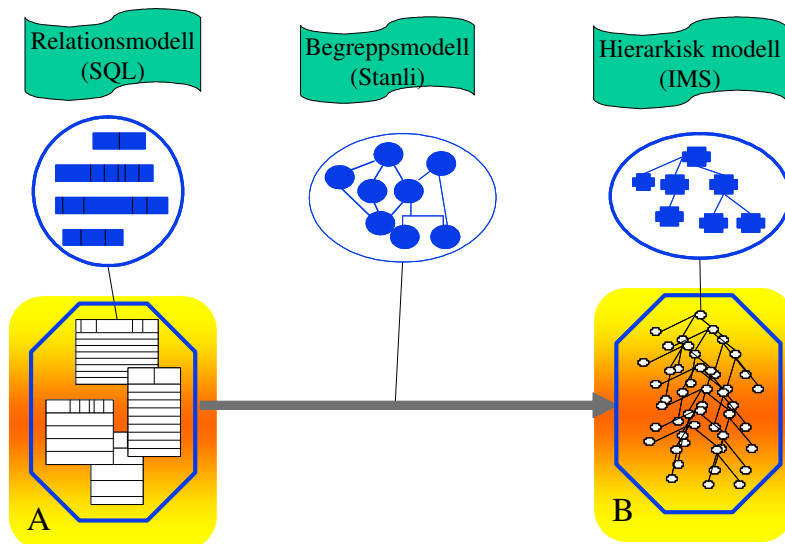
Tillbaka till A och B. De har kommit fram till en gemensam begreppsmodell som fungerar för de aktuella behoven. Det kan vara en modell beskriven i Stanli-språket (se separat rapport), UML ..... Alternativen är som sagt många. Det viktiga är att modellen passar A och B. Anta att data ligger lagrade hos A och B i olika tekniklösningar och tillämpningar som dessutom representerar unika begreppsmodeller för varje intressent.



Figur 4 a

Figur 4 b

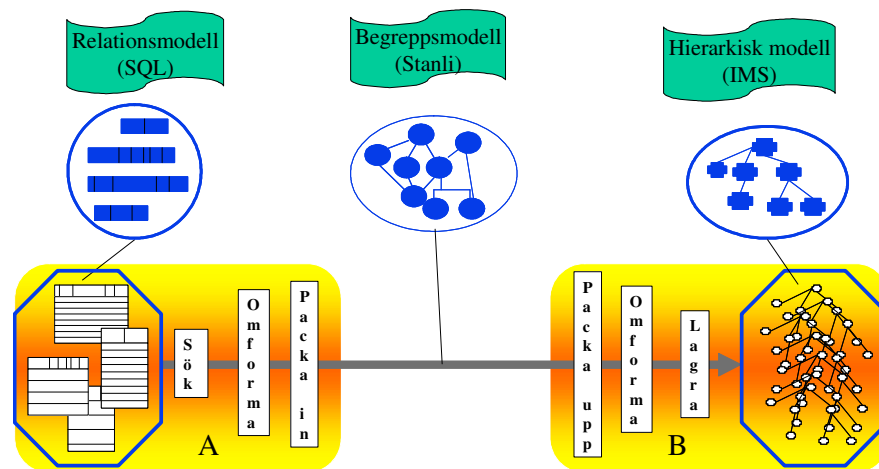
Vad man är överens om - och behöver vara överens om - är den begreppsmodell som ska gälla vid själva överföringen av information. Vi antar här att denna begreppsmodell är uttryckt i Stanli-notation.



Figur 5

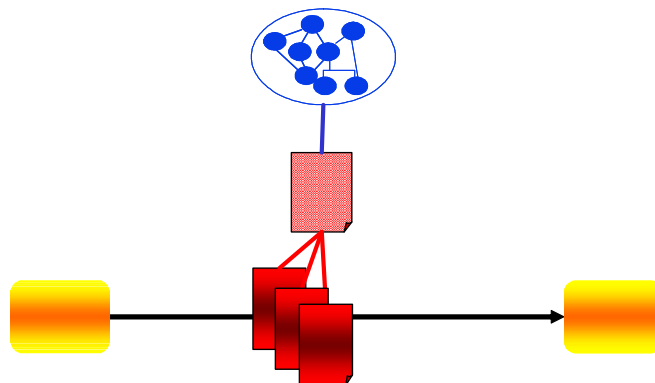
För att ansatsen ska fungera i praktiken behöver ett antal saker utföras både hos A och B. Anta att A är leverantör och B mottagare av information. Den relevanta informationen måste först sökas i As interna kunskapsbank/databas med hjälp av det utsökningsbehov som formulerats och i det gränssnitt som står till buds. Alternativt sker kontakt med någon tillämpning som erbjuder ett lämpligt gränssnitt.

Därefter måste informationen föras över i den form som den gemensamma begreppsmodellen kräver. Det kan gälla översättning av begreppstermer, ändring av informationens struktur, mm. Därefter måste den omformulerade informationen packas in i ett överenskommet format för själva överföringen. Här kan XML eller annan syntax tänkas bli använd. På Bs sida behöver informationen tas om hand med i princip motsvarande steg, dock i omvänd ordning. Det är knappast en överdrift att påstå att risken för fel och inkonsistenser är överhängande med denna ansats.



Figur 6

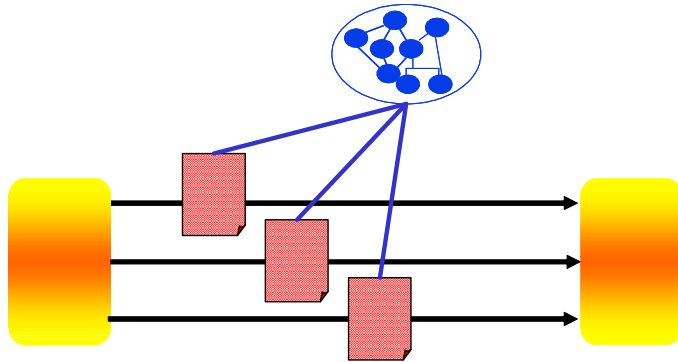
Hur samspelet mellan begreppsmodeller och meddelandetyper är tänkt att utformas framgår mer i detalj i figur 7.



Figur 7

Informationen i ett XML-meddelande (dokument) struktureras i och för överföring i enlighet med en definierad meddelandetyper (DTD eller schema). Meddelandetyper beskriver betydelsen – semantiken – hos varje ingående elementtyp med referens till den förenande begreppsmodellen. Dessa referenser kan i det enklaste fallet vara implicita såtillvida att märkorden har sin motsvarighet i begreppsmodellen. I en mer ambitiös lösning finns såväl begreppsmodell som meddelandetyper definierade med sina explicita relateringar i en och samma beskrivningsdatabas.

Finessen ligger primärt i det förhållandet att samtliga meddelandetyper refererar till samma begreppsmodell. Samordning, gemensamgörande, konsistens åstadkoms ”med automatik”.



Figur 8

Vi får i sammanhanget inte heller glömma problemet med att omformulera en begäran om utsökning i de fall den sker spontant och inte är fördefinierad eller integrerad med en tillämpning. För As och Bs samverkan gäller det att översätta från en fråga uttryckt i enlighet med villkoren för det hierarkiska språket, över till en formulering anpassad till, Stanlis krav (i exemplet) och vidare över till SQL.

Observera att de beskrivna förutsättningarna knappast kan anses unika eller onödigtvis konstruerade, i alla händelser inte när det är fråga om många huvudsakligen fristående, samverkande intressenter. Samverkan ställer krav.

## 2.4 XML Schema som modelleringsspråk

XML Schema är i första hand avsett som ett språk för att definiera meddelandetyper, att erbjuda bättre uttrycks kraft i olika avseenden än vad XML 1.0 kan erbjuda. XML Schema kom - genom ett antal utvecklingsturer - att hämta intryck från ”vanliga” modelleringsspråk vilket gav till resultat att man så småningom - som en ”bonus” - visade sig kunna använda språket även för informationsmodellering. I alla händelser i viss utsträckning.

Hur man än vänder och vrider på det är XML Schema för informationsmodellering uttrycksmässigt en efterrapning - ett uppfinna hjulet igen - resultat. Varför man inte accepterade att UML med flera språk redan finns väl uttestade sedan några decennier tillbaka för informationsmodellering och valde att koncentrera XML Schema som ett språk just för meddelandetyppspecifikation, är en gåta. Känns som ett så onödigt slöseri med resurser plus att det bara tillför oro och komplikationer när dokument- och informationsvärldarna ska mötas. Förmodligen såg man inte i arbetets inledningsskede någon distinktion mellan informationsmodeller och meddelandetyper alternativt inte alls behov av stödjande informationsmodeller. Möjligtvis är detta för en hel del fortfarande ett faktum.

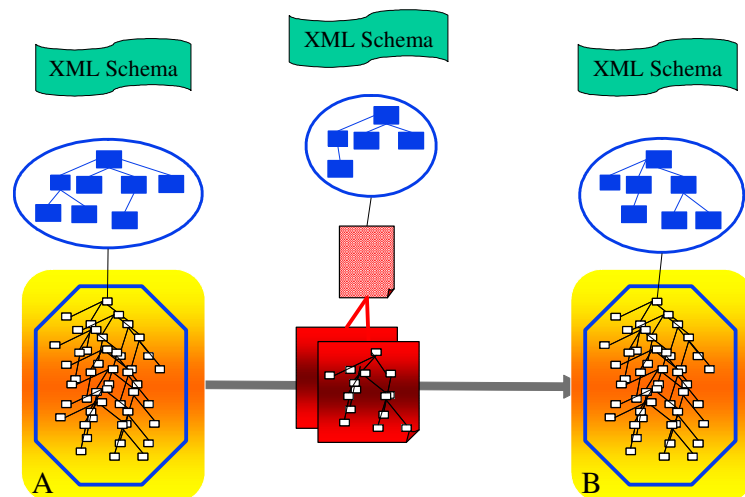
I en separat utredning hoppas vi få möjlighet att titta närmare på för- och nackdelar i det avseendet. Nåväl, med facit i hand, kan alltså XML Schema, åtminstone i vissa avseenden,

användas för att uttrycka den generella utbytesmodellen i figur 8 ovan, förutom som specifikt meddelandetyppspråk. Referenser från meddelandetyper till utbytesmodellen kan exempelvis göras via Namespaces-specifikationer om man fullt ut vill hålla sig till XML-miljön. Namespaces erbjuder dock inga explicita relateringar. Den principiella skillnaden mellan att referera till en begreppsmodell uttryckt i XML Schema och en annan uttryckt i exempelvis UML är möjligtvis inte överhängande. Vad som däremot är viktigt är en godtagbar uttryckskraft i det språk som används för begreppsmodellen i kombination med den övergripande strategin att använda en begreppsmodell som semantisk referenspunkt.

### 3 Samma modelleringspråk "över hela linjen"

Anta nu den unikt smakliga situationen att A och B båda opererar med samma typ av intern databaslösning. Anta än bättre den i realiteten osannolika omständigheten att denna typ av databashanterare hos båda hanterar XML-dokumentsdatabaser. Sådana finns även om de har lite varierande faciliteter. Låt oss för enkelhets skull ytterligare anta att exakt samma databashanterare (i samma version) används av båda och att den är så modern att den definierar sina datastrukturer med hjälp av XML Schema.

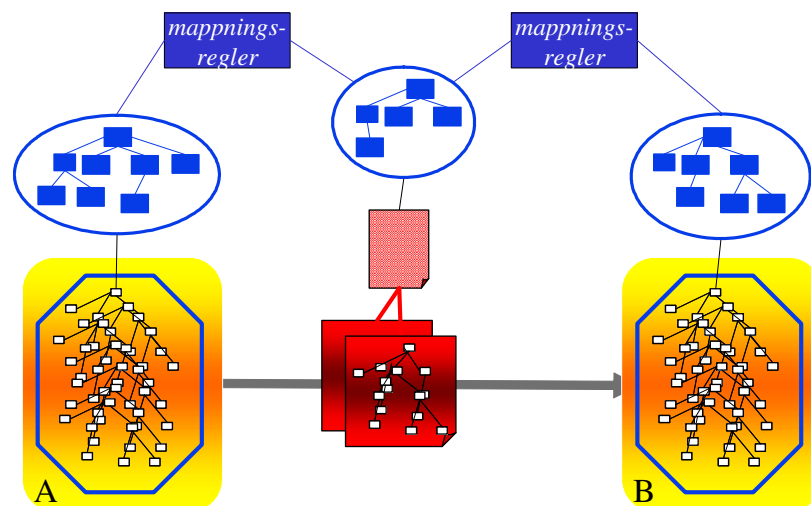
Därmed är väl i princip alla bekymmer med samordning, transformation, mm ur världen? Det är bara att skyffla önskad information fram och tillbaka på enklaste sätt?



Figur 9

Ett steg i förenklande riktning är det nog, speciellt vad gäller diverse teknikrelaterade problem. Dock hjälper knappast enklare tekniska förutsättningar om A och B inte internt delar samma syn på den aktuella informationsvärlden. Det hjälper knappast att respektive uppfattning är formulerad i samma modelleringspråk. Olika begrepp, taxonomier och strukturer kan mycket väl existera hos respektive intressent även om de delvis "tittar på" samma verklighet. Fortfarande behövs omformuleringar från den interna till den gemensamma överföringsmodellen för A och tvärtom för B.

Förhoppningsvis är denna omformulering nu ändå något enklare eftersom samma modelleringspråk används för det som ska omformas och det omformade. Möjligtvis kan till och med ett standardiserat språk för omformuleringar (mappingsregler) definieras så att arbetet kan automatiseras utifrån specifikationer gjorda i språket. Sannolikt blir även in- och upppackning enklare av samma skäl. Datatyperna är kompatibla. Varje meddelande blir ett XML-dokument formulerat i enlighet med XMLs regler. Innehåll är ett utsnitt ur den levererande databasen. Dock kvarstår som sagt kravet på att detta utsnitt strukturellt måste anpassas till och från utbytesmodellens villkor. I förlängningen är det inte omöjligt att något finurligt mappningsstöd skulle kunna automatisera mappningarna, givet att reglerna kunde specificeras som relateringar mellan respektive intern modell och den generella.

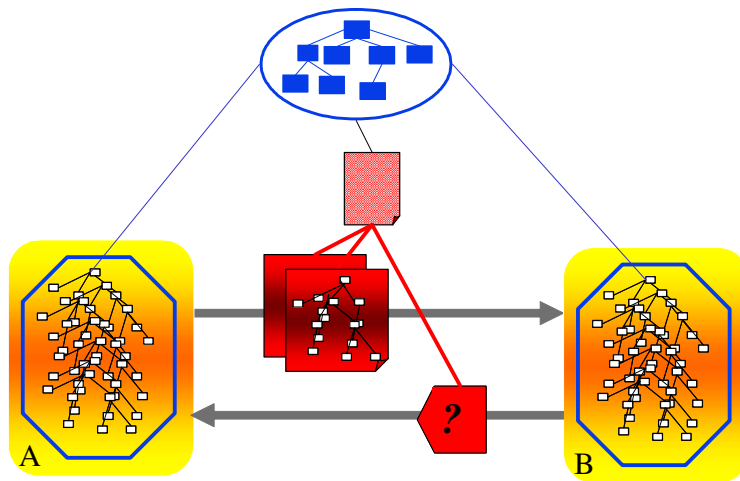


Figur 10

### 3.1 Gemensam begreppsmodell

Ytterligare ett steg mot en ideal värld är en situation där A och B enligt ovan inte bara använder samma modelleringspråk samt har lyckats formulera den gemensamma begreppsmodellen i och för överföringar. Nu visar det sig att de i grunden har samma uppfattning om verkligheten, artikulerat genom att de nyttjar samma begreppsmodell för såväl den interna hanteringen som för utbytet. Med intern hantering menas inte nödvändigtvis att informationen ligger lagrad på ett visst sätt, bara att det finns mekanismer för att åskådliggöra och exportera den interna informationen i enlighet med angiven begreppsmodell.

Även om de inte har totalt samma uppfattning om verkligheten har de kanske valt att tillämpa en standardiserad begreppsmodell. Det är praktiskt och underlättar informationsutbyte inom den egna "branschen" eftersom begreppsmodellen är allmänt spridd. Samsynen är total. Inga omformuleringar behöver tillgripas. Förmodligen finns teknikstöd anpassade till begreppsmodellen som mer eller mindre automatiserar in- och upppackning. Förmodligen utnyttjas dessutom ett standardiserat språk för formulering av utsökningsbehov. Komplicerade översättningar av dessa kan undvikas.



Figur 11

Låter inte detta alltför bra för att vara sant? Kanske inte inom branscher där intressenterna känner varandra och arbetar med samma grundsyn. Inte heller mellan andra partsgrupperingar som sökt kontakt med varandra för informationssamverkan på grund av trängande behov, av pur bekvämlighet eller någon form av värdegemenskap.

I en samverkansmiljö med många intressenter och där miljön inte etablerats som en grundförutsättning för den egna verksamheten utan som ett mer perifert tillskott, tillstöter säkert komplikationer "som ett brev på posten".

## 4 Semantik och informationskvalitet

Trots existensen av en gemensam begreppsmodell som uppstyrande facilitet kvarstår alltså förmodligen en del potentiella bekymmer att ta tag i. Låt oss diskutera några av dessa utifrån omständigheterna i figur 11. Är det problem där är det garanterat minst lika stora problem för förutsättningarna i de tidigare figurerna.

Anta att den av A och B anammade begreppsmodellen är branschens egen verksamhetsmodell VM. A har dock valt att inte fullt ut nyttja alla de attribut som VM erbjuder. Samma sak gäller B. Så länge respektive lokal del av VM är en delmängd av den gemensamma samverkansmodellen är situationen rimligt hanterbar. Dock kan bekymmer ändå uppträda. Kanske använder sig A och B av delvis olika taxonomier för ett visst attribut. Kanske har B kvalificerat vissa attribut på ett sätt som A inte valt att göra eller valt att göra på annat sätt. Hur ska B kunna föra in As uppgifter under sina kvalificerade attribut om inte A alls levererar med kvalificerade eller levererar med andra sådana?

Anta att B helst skulle vilja ha lite mer information än A kan erbjuda. Förmodligen blir den gemensamma begreppsmodellen anpassad till vad A klarar av. As lokala modell är en delmängd av den gemensamma som i sin tur är en delmängd av Bs. B kan ta emot alla data, men hur ska i ett senare läge de tomrum som As ofullständigheter (ur Bs perspektiv) orsakar tolkas?

Ta informationsutbyte kring litteratur. Är utebliven uppgift om 'Förläggare' att tolka som att boken inte givits ut, att förläggaren är okänd, att förläggarpuppgift saknas dvs. att förläggare finns men inte kunnat identifieras, att ....?

Anta att A endast dokumenterar den förläggare som ger ut förstaupplagan och att B väljer att dokumentera samtliga förläggare. Anta vidare att A översänder en förläggare till en bok till B och B dokumenterar uppgiften hemmavid. Vid ett senare tillfälle söks förläggarinformation ur Bs kunskapsbank. Frågeställaren får fram vissa böcker med flera förläggare, andra med bara en, bland annat förläggaren till As bok. Frågeställaren utgår ifrån att svaren är representativa enligt Bs begreppsmodell, d.v.s. att om det finns flera förläggare kommer samtliga att redovisas. När det gäller As bok visas endast förstautgåveförläggaren. Kanske är denna uppgift det fullständiga svaret, kanske inte. Finns flera förläggare till boken blir säkert inte frågeställaren glad över att bara få information om en av dem. Kvaliteten i Bs bank är inte fullgod - som en följd av informationsutbytet med A.

I grund och botten har inte B en aning om kvaliteten i den information som A skickar enbart utifrån den gemensamma begreppsmodellen. Om inte kvalitetskraven reglerats på annat sätt förstås. Varje attribut i VM är ju frivilligt och oberoende av övriga.

Över till en annan omständighet. Det är inte osannolikt att innehållet i såväl As som Bs kunskapsbank skapats över tiden. I början orkade man inte mata in så mycket data per bok. Allteftersom har man blivit duktigare eller blivit utsatta för hårdare användarkrav. De tidiga uppgifterna är utifrån ett senare perspektiv alltså rejält ofullständiga, har bristande kvalitet sett som helhetsbeskrivning av boken. Hur handskas man internt och inte minst i informationssamverkan med denna förutsättning? I vissa fall kan man säkert "leva med" ofullständigheter, i andra sammanhang kan det få kritiska effekter.

Sannolikt kan många fler exempel på bekymmersamma omständigheter av motsvarande slag målas upp. Detta trots att förutsättningarna bygger på en gemensam begreppsmodell.

## 5 Gemensamgörande i praktiken

Erfarenhetsredovisningar från internationella projekt pekar på att förutsättningar enligt figur 11 sällan eller aldrig uppstår i realiteten. Inte minst är detta ett ofta konstaterat faktum om begreppsmodellen tagits fram genom ett internationellt standardiseringsarbete för att svara mot behov inom visst tillämpningsområde. Det beror inte så mycket på konflikterande krav mellan intressenterna som på att begreppsmodeller inte täcker in samtliga gemensamma krav - inte har tillräcklig begreppsrepertoar eller företräder betydelser av begrepp som inte överensstämmer med intressenternas uppfattning. Standarden som kompromiss kan ha blivit "urvattnad", alternativt i huvudsak speglade viss dominerande intressents intressen alternativt inte orkat anpassa sig till aktuella omständigheter.

I dessa lägen väljer intressentgruppen att komplettera och revidera. Kanske väljs vissa begrepp från en standard, andra från en konkurrerande standard. Kanske tillför man egna när ingen standard erbjuder alternativ. Resultatet blir visserligen en fungerande begreppsmodell intressenterna emellan men sluter samtidigt intressentkretsen från omgivningen. Man brukar i

sammanhanget tala om 'application profiles'. Ibland är det helt acceptabelt, i andra fall förödande. Men vad är alternativet? Att plocka in så mycket som möjligt av repertoaren i standardiserade begreppsmodeller är bättre än att hitta på allt på egen hand. De data som svarar mot en standard kan ju i alla händelser (om vi bortser från problemresonemangen ovan) utbytas med sådana andra intressenter som företräder samma standard.

Någon genväg till lösning finns inte. Erfarenhetsredovisningar menar på att endast ett ansvarsfullt pragmatiskt arbete står till buds både när det gäller att forma begreppsmodeller för egna behov och för gemensamma. Att diktera lösningar håller aldrig i längden. Orsaken till olika krav på begreppsmodeller beror nämligen mycket sällan på allmän arrogans och illvilja hos intressenten utan på välgrundade skillnader i uppfattningar och behov. Erbjuds inte möjlighet att forma modellen efter de egna behoven inskränks också nyttan med informationshanteringen och informationsutbytet. Att hoppas på en central modell för alla behov är alltså sällan realistiskt, i alla händelser inte om och när ambitionsnivån för samverkan expanderar till mer än bara de mest rudimentära elementära begreppen i någon standard.

Vi är i grunden tillbaka till de samverkansförutsättningar som figur 10 skissar. Många intressenter tillgodoser delbehov av samverkan genom upprättande av en begreppsmodell för respektive behov, för respektive roll de spelar i olika sammanhang. Sker det med största möjliga anpassning till vad begreppsmodellstandarder erbjuder undviks i görligaste mån onödiga inlåsningar.

Sannolikt överskuggar de semantiska problemen vid informationsutbyte i de flesta fall vida de tekniska bekymren. Tillsammans är de ett problemområde som borde generera betydligt mer intresse inom SOA än hittills varit fallet.