



Subproject: *Supporting facilities*

Tema Informationssamverkan - Rapportserieöversikt¹

**Version 0.5
(2004-05-05)**

Stig Berild

(Santa Anna IT Research Institute AB)

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

1 Inledning

Föreliggande översikt introducerar ett antal rapporter och noteringar som skrivits, skrivs och kommer att skrivas under ett aktuellt samarbete mellan projektet e-Society II (www.skriver.nu/esociety) vid Santa Anna IT Research Institute AB (www.santaanna.se) och Serviam-projektet (www.serviam.se). Därutöver refereras ett antal rapporter som skrivits under år 2003 med stöd av Myndigheten för skolutveckling.

Gemensamt för rapporterna är fokus på informationssamverkan med relatering till de nya intressanta arkitekturer som brukar gå under beteckningen Service Oriented Architecture (SOA).

Vissa rapporter är grundligt genomarbetade medan andra är lösare skisser och funderingar.

Samarbetet beräknas pågå under de två år (2004 - 2005) som är Serviams aktuella projekttid men med de huvudsakliga insatserna förlagda till det första projektåret. Insatserna från e-Society II finansieras av Stiftelsen SISU. Serviam är ett projekt drivet av Dataföreningen och med stöd från Vinnova. I Serviam deltar dessutom ett tiotal företag med såväl finansieringsbidrag som egna insatser.

Översikten syftar till att placera in de olika rapporterna i ett sammanhang och att påvisa deras tangeringspunkter. På så vis kan det introducerande avsnittet i varje rapport hållas till ett minimum. I själva verket finns i flera av dem endast en referens till den aktuella översikten.

Av förklarliga skäl bör översikten betraktas som ett arbetspapper under ständig revidering så länge som samverkan pågår, nya rapporter kommer fram och kunskaperna inom området expanderar och formerar sig. Av samma anledning har översikten alltid ett aktuellt versionsnummer i kombination med ett datum.

Publicerade rapporter, oavsett om de är gediget genomarbetade eller lösa funderingar, noteras i referenser till litteraturlistan med fet stil. Dessa rapporter finns för nerladdning på webbplatsen www.skriver.nu/esociety. Medlemmar i Serviamprojektet kan även ladda ner rapporterna från www.serviam.se.

Planerade rapporter noteras kursiverade. Planerad kan innebära allt från en senare publicering till strykning i nästa version av översikten.

Låt oss börja.

2 Aktuell trend

Vi befinner oss bara i början av en intressant utvecklingstrend kring webben och dess potential. Diskussionerna kring Service Oriented Architectures (SOA) och där kanske speciellt kring den nya Web Services-arkitekturen pågår intensivt med alla större ICT-aktörer inblandade. [1]

Ett ökat verksamhetsstödande fokus är påtagligt samtidigt som teknikoberoende eftersträvas. I visionen ligger förhoppningar om att tjänster ska kunna sökas, hittas, avtalas och tas i anspråk mer eller mindre utan mänsklig inblandning. Tjänster ska kunna utbyta adekvat information baserad på en gemensam semantisk förståelse i form av gemensamt överenskomna begreppsmodeller. Även här finns strävan att minimera den mänskliga inblandningen. Kontakter/samarbeten kan vara temporära eller mer permanenta, bygga på övertänkta planer eller på impulser, vara raka och enkla eller komplexa med många parter involverade, kräva avtal eller ”gå på chans”, osv. Dynamik i högsätet – en dynamisk webb.

Klyftan mellan denna vision och något realiserbart behöver inte vara avgrundsdjup. I alla händelser inte om ambitionen är att endast ta ett steg på vägen. Förutom att beskriva tjänster med de operationer de är kapabla att utföra, skulle en kompletterande beskrivning av allmänna kapaciteter, ansvar, aktuella arbetsuppgifter, drivkrafter, roller samt kontaktytor mot omvärlden kunna lägga grunden till än mer intressanta samarbetsmönster.

Här återstår dock fortfarande mycket arbete. Visserligen har standarder tagits fram för att beskriva, publicera och söka information om tjänster. Bland dessa standarder återfinns till exempel UDDI och WSDL. De fungerar säkert utmärkt för sina ändamål. Vad som kvarstår, och i förlängningen kan bli betydligt allvarigare, är det ringa intresset för vad tjänst kan tänkas vara för slags företeelse, det ringa intresset för att definiera de olika betydelser som kan gömma sig bakom termen ”tjänst”. Utan klarare begreppsbildning riskerar intressenter inom området att tala förbi varandra med de effekter av negativ art dylikt alltid för med sig. [2]

En strävan inom området är att försöka vända intresset mot så kallade verksamhetsstödande tjänster för att med detta påtala att tjänster och tjänstesamverkan inte alls enbart är en företeelse för teknikskapare att intressera sig för. De nya arkitekturerna syftar till att även erbjuda olika verksamheter samarbetsmöjligheter. E-handel, B2B samverkan, mm. har banat vägen. Tyvärr råder även stor osäkerhet om vad som utmärker något som ska anses vara ”verksamhetsstödande”, kanske till skillnad från något ”teknikstödande”. [3] [4]

Tjänster står alltså i centrum för det dagsaktuella intresset. Betydligt mildare intresse riktas mot den information som i de allra flesta fall formulerar förutsättningarna för tjänsternas möjliga agerande. Informationen är ju trots allt ofta det livgivande blodet i kommunikationsådorna. Det svala intresset är något förvånande med tanke på att det var informationsutbyte (med hjälp av XML), som utgjorde de initiala ingredienserna i Web Services Architecture. De inledande bristerna hos XML har dock åtgärdats genom kompletterande standarder. Den kanske viktigaste i det avseendet är XML Schema, den standard som förväntas bli den dominerande för definition av strukturer och format i de meddelandetyper som reglerar XML-dokuments innehåll. [5]

Men information, informationshantering och informationsutbyte är mycket mer än bara XML.

Kanske anser man sig nöjd med de förutsättningar som XML och XML Schema erbjuder. Kanske erbjuder tjänsteperspektivet mer av utmanande teknikrelaterade problemställningar. Kanske har de ledande programvaruleverantörerna konstaterat att potentialen för att tjäna pengar ligger just i att erbjuda teknikstöd för arkitekturerna. Information, informationshantering och informationsutbyte är genom sin tillämpningsorientering betydligt mer svårfångad och betydligt mindre generell till sin karaktär. Bland tekniker har heller aldrig information givits en speciell framträdande plats.

Häri ligger en definitiv fara. Riskerna är att vi alla lockas in i en tro att bara arkitektur och tjänster är på plats är den nya härliga samverkansvärlden realiserad och klar och målet nått. En i högsta grad riskfylld inställning. Inom väl avgränsade tillämpningsområden involverande en begränsad mängd intressenter är förmodligen både risker och problem av hanterbar natur. Men så snart vi närmar oss de nya trendernas vision om öppenhet, flexibilitet, anpassbarhet blir problembilden genast betydligt mer komplicerad. Att den sällan kommer upp på agendan beror förmodligen på att de flesta realiserade tillämpningsmiljöer fortfarande är relativt avgränsat hanterbara. Problemen har ännu inte nått verkligheten. Men tidsfristen är begränsad. Alltså gäller det att också fånga de problemkomplex som omger information i en öppen samverkansarkitektur, dels om möjligt formulera strategier för effektiva lösningar.

3 Informationsutbyte med traditionella förtecken

Behov av informationsutbyte mellan intressenter noteras och växer vid mer etablerade förutsättningar fram genom konventionell verksamhetsanalys, genomgång av användningsfall, scenarier, mm (förutom genom erfarenheter från redan driftsatta system). Traditionellt har databaser utgjort en typisk arena för informationsutbyte inom väl avgränsade tillämpningsområden, med väl identifierade aktörer och deras olika roller. Databastillämpningen representerar en integrerad enhet. Likafullt får vi inte glömma att även denna miljö erbjuder ett ofta synnerligen avancerat informationsutbyte. [6]

Utbyte under friare former och mellan mer självständiga aktörer realiserar snarare med hjälp av meddelandeutväxling. Ett antal meddelandetyper specificeras och överenskomms aktörerna emellan. Varje aktör ser för egen räkning till att de överenskomna meddelandetyperna kan hanteras.

På avsändarsidan gäller det sökning efter eller konstruktion av aktuella informationselement, inpackning av desamma samt leverans över ett eller flera medier. På mottagarsidan gäller på motsvarande sätt uppackning av informationselementen samt distribution till avsedd mottagare för operation, lagring eller dylikt.

Inte sällan underlättas arbetet genom att branschorganisationer, de facto standardiseringsorgan och andra formulerar meddelandetyper för rekommenderad användning inom givna tillämpningsområden.

Många olika alternativ står till buds för den fysiska paketeringen av informationen i och för transport. EDIFACT-standarderna är exempel på en sedan länge väl etablerad standard för såväl innehåll som syntax. XML är ett annat exempel som, speciellt inom SOA, etablerat sig som

det i princip enda språket för formulering av meddelanden och meddelandetyper. XML bedöms kunna tjänstgöra som förpackningssyntax för de flesta utbytesbehov.

Vid stabila förutsättningar inom väl etablerade tillämpningsområden fungerar ansatsen förmodligen utmärkt. Är intressenterna dessutom väl förtrogna med varandra, kanske tillhörande samma koncern eller någon branschintern gruppering är läget än bättre. Meddelandetyperna har sitt rimligt fasta innehåll. Uppdateringar sker relativt sällan samt, när det väl sker, genom väl etablerade rutiner.

Nya meddelandetyper införs sparsamt eftersom varje sådan meddelandetyper måste diskuteras och genomlysas från grunden utifrån sina villkor och förutsättningar. Varje informationselement och dess relatering till övriga informationselement inom meddelandetyperna måste analyseras såväl till semantisk innebörd som till de övriga villkor och frihetsgrader som ska gälla för acceptabel formulering av varje individuellt meddelande.

Dessutom kräver förmodligen uppdateringsarbetet god kunskap om redan etablerade meddelandetyper, deras innehåll och struktur samt om de resonemang som legat bakom den resulterande utformningen.

Varför då? Jo, därför att meddelandetyperna tillsammans antagligen svarar mot samma tillämpningsområde och därför med stor sannolikhet har tangerande och kanske delvis överlappande innehåll. Om och där så är fallet gäller det att se till att sådant överlappande innehåll får samma utformning (semantik, villkor, struktur) i varje meddelandetyper, att informationen formuleras på samma sätt oavsett vilken meddelandetyper samt används för leverans av i grunden samma information.

Vilket osökt för oss in på begreppsmodeller.

4 Begreppsmodeller som fundament

Även inom givna verksamhetsområden kommer företag och organisationer att i allt större utsträckning verka under dynamiska förutsättningar. Man vet i princip sin egen roll i sammanhanget. Men nya samverkansparter kommer, andra försvinner. Kanske ställs krav på anpassning till dessa i form av nya informationselement eller i form av nya eller modifierade meddelandetyper. Verksamhetsområdet utvecklas. Nya informationsbehov uppstår, existerande anpassas efter nya förutsättningar eller försvinner. Och så vidare.

Att under dessa förutsättningar basera informationsutbyte på ett antal fristående meddelandetyper blir i längden tungt och osäkert. Det behövs någon form av fundament för dynamiken att vila på. In på arenan kommer begreppsmodeller.

En alldeles avgörande förutsättning för meningsfull samverkan, oavsett om den sker inom en liten eller stor grupp, om den sker lokalt eller globalt, är möjligheten för de samverkande intressenterna att kunna uppfatta och förstå den information som utbyts, att kunna tolka betydelsen hos de termer som utbyts. Människor emellan är kanske inte alltid exakt tolkbarhet ett krav. Vi har olika medel för att fånga och hantera otydligheter. Digitala tjänster emellan är kravet på exakt tolkbarhet i de allra flesta fall betydligt större. Tolkbarhet underlättas genom tillgång till en mellan de samverkande intressenterna överenskommen begreppsmodell.

(Observera att märkorden i XML-dokument knappast kan sägas förmedla någon semantisk betydelse.)

Begreppsmodeller och informationsmodeller är ingenting nytt. De har i ett antal decennier spelat centrala roller för att stadga upp den semantiska dimensionen och stringensen i olika verksamheter men kanske framförallt vid strukturering och drift av databastillämpningar.

UML (Unified Modeling Language) är förmodligen det mest kända och spridda språket för objektmodellering. För begreppsmodellering anses Stanli-språket [7] fungera utmärkt. Detta språk utvecklades inom Stanli-projektet (www.stanli.se) i början på 1990-talet. Stanli ingår numer i SIS (www.sis.se).

Även så kallade binära språk har vunnit i popularitet sedan W3C nyligen antog RDF-standarderna. Binära språk är dock ingenting nytt. De första specificerades för mer än tre decennier sedan och har därefter kommit till användning både för modellering och som strukturingsmetod i databashanterare. [8]

Begreppsmodellens nya användningsområde som "semantisk mötesplats" för självständiga parter med olika geografisk hemvist, olika sociala och kulturella traditioner, mm ställer nya krav på dessa språk, krav som de i dagsläget inte svarar upp emot. Bland dessa är möjligheten till distinkt åtskillnad mellan begrepp och term så att olika termer valfritt kan användas för att representera samma begrepp. Det gäller alltså att kunna definiera mappningar mellan olika termers symbolvärldar, till exempel så att oavsett om man anger "1", "Female", "Kvinna" eller "<en bild på ett kvinnoansikte>" så står det för samma valör. Funderingar kring möjliga språkutvidgningar för att klara dessa behov finns dokumenterade i [9] och [10].

Den som är intresserad av en snabbintroduktion till några aspekter att tänka på i största allmänhet vid begreppsmodellering hänvisas till [11].

5 Begreppsmodellstött informationsutbyte

Traditionellt har varje tillämpning utgått ifrån sin egen "lokala" begreppsmodell. Samtidigt kan konstateras att dessa mer eller mindre "isolerade öar" i allt högre grad behöver kunna samverka i och för informationsutbyte för olika syften. De är i den rollen intressenter kring extern information. Samverkansbehov mellan intressenter föreligger naturligt lokalt inom enskilda verksamheter. I allt ökande utsträckning behöver dessa verksamheter med tillhörande tillämpningar även knyta externa kontakter mellan etablerade affärspartners.

Därutöver kommer samverkan av mer tillfällig och spontan natur att bli allt vanligare mellan såväl kända som okända intressenter. I samtliga fall gäller att geografiska och andra hinder ska vara undanröjda. Det ligger i tiden att tala om det globala perspektivet, om cyberspace.

Samverkande intressenter har ett gemensamt intresse av att utbyta information inom visst tillämpningsområde eller för visst syfte. Det gemensamma intresset formuleras i form av en gemensamt överenskommen begreppsmodell enligt argumentationen ovan.

Begreppsmodellen är relativt stabil eftersom den representerar en modellbild av den verkliga värld som av någon anledning är av intresse att utbyta information om. Begreppsmodellen ger en neutral helhetsbild.

Däremot kan man förvänta sig att intressentrelaterade specifika behov efterfrågar olika skärningar av den totala informationsmassan. Dessa skärningar kan vara allt från ett enkelt informationselement, över mer samma meddelandetyper till omfattande vyer av helheten. Nya eller förändrade utbytesbehov eller meddelandetyper uppstår över tiden, mer eller mindre frekvent.

Det nya, i de ansatser som rapportserien för fram, i förhållande till konventionellt informationsutbyte, är bindningen av varje meddelandetyper till begreppen i begreppsmodellen. Härigenom definierar varje meddelandetyper sitt innehåll med referens till ett gemensamt fundamentet, nämligen den överenskomna begreppsmodellen. Begreppsmodellen blir den gemensamma semantiska nämnaren för entydig tolkbarhet.

Jämför med databasscheman där man med tillgängliga frågespråk kan formulera valfritt många olika frågetyper så länge som de har en tolkbar innebörd gentemot schemat. [6]

Eftersom varje meddelandetyper i den här ansatsen definierar sitt innehåll med hjälp av referens till begreppsmodellen bortfaller dessutom risken enligt tidigare resonemang för inkompatibla definitioner av i princip samma informationsinnehåll i olika meddelandetyper.

Skulle en uppdatering av begreppsmodellen behövas, framgår det därifrån indirekt med automatik vilka meddelandetyper som eventuellt påverkas och i sin tur måste uppdateras. Självfallet har en begreppsmodell, liksom allt annat som lever i direkt relation med en föränderlig värld, en livscykel som måste kunna hanteras. Cykeln kan vara kort om begreppsmodellen endast har att svara mot vissa temporära behov. Sannolikare är förmodligen att den representerar ett samverkansbehov av mer permanent art.

Varje tillämpningsområde och konstellation av samverkansparter definierar sina behov i form av en egen begreppsmodell. Tillämpningsområden kan vara omfattande men lika väl synnerligen avgränsade. Konstellationen kan vara allt ifrån ett internationellt standardiseringsorgan, över en branschorganisation till den lokala företagarföreningen.

Notera att begreppsmodellen endast definierar den gemensamma synen på information i och för utbyte. Vad varje intressent internt tycker "hemma hos sig" är en ren privatsak.

Sedan är det inte osannolikt att begreppsmodellen mycket väl - i den mån den visar sig effektiv - successivt blir en alltmer accepterad begreppsbas som även de medverkande intressenterna successivt väljer att ta till sig för den egna interna miljön. Dels är det praktiskt, dels undviks transformeringar till och från egna modeller vid utbyten.

Inte minst kan det vara så att begreppen helt enkelt blir alltmer etablerade och accepterade - känns rätt. En väl fungerande begreppsmodell kan under sin livscykel ges utökade revir till sitt innehåll men likväl expandera sin acceptans till allt bredare lager.

Olika principer för användning av begreppsmodellstött informationsutbyte diskuteras i [12].

En friare form av informationssamverkan, som bygger på parternas egenexponering, av den typ av information som man är villig att dela med sig av, diskuteras i [13]. Denna exponeringsinformation måste också struktureras, samordnas, standardiseras så att den blir opererbar på ett standardiserat sätt oavsett vem som exponerar vad. Vilket öppnar för att många kan vända sig till många med samma frågeställning.

Därmed är vi inne på vad som normalt kallas metadata.

6 Metadata

Metadata brukar stå för data om data. Ofta åsyftar man de data som används för att beskriva digitala resurser, gärna resurser (musik, video, rapporter, ...) som finns att tillgå på webben. De katalogsystem som traditionellt används inom biblioteksvärlden för att beskriva och söka litteratur är ett annat exempel. Kanske är det just härifrån och från dokumenthantering som de initiala erfarenheterna har hämtats och vars traditioner hittills kommit att influera synen på och hanteringen av metadata även vid introduktionen i webbmiljön.

Detta är ett potentiellt bekymmer eftersom man i de influerande världarna har en delvis annan tradition, när det gäller hur dessa typer av data ska struktureras och representeras, än vad som finns företrädd i exempelvis konventionell informationshantering. Rapporten [14] argumenterar för att de upplevda skillnaderna mellan data och metadata med därtill hörande olika angreppssätt och olika struktureringsspråk saknar grund samt föreslår en utväg ur dilemmat. Om inte blir utfallet endast onödiga otydligheter, missuppfattningar och försvårat informationsutbyte. Samma rapport finns på engelska. [15]

Ett annat bekymmer som förtjänar en fördjupad belysning är att betydelsen av termen "metadata" långt ifrån är kristallklar. Något som [16] helt kort noterar.

7 Binär arkitektur

Idén med begreppsmodellstött informationsutbyte inom SOA leder naturligt tankarna vidare mot en mer renodlad arkitektur för ändamålet än vad Web Services förmår prestera. Speciellt gäller detta i kombination med de trender som kan skönjas med relatering till exponering (som diskuteras i [13]). Läg därtill de tankegångar som finns företrädda inom semantisk webb-sfären och spännande utmaningar öppnar sig för vidare exploatering.

En ansats till en mer renodlad arkitektur finns beskriven i rapporterna [17a – 17n].

8 Dynamisk webbmiljö

Anta nu att inte bara information om webbresurser och informationsmodeller beskrivs som metadata tillgängligt på webben. Komplettera med tjänster och deras meddelandetyper in och ut, med samverkansvillkor samt diverse annan intressant information om respektive part. Här finns en rik källa av uppgifter för att kunna bedöma en part, etablera kontakt med parten och därefter enligt överenskomna villkor att utbyta information, anlita tjänster med mera.

Anta dessutom att denna information finns beskriven på ett sätt som både är rimligt fullödigt för att vara av intresse och har en sådan semantisk precision i beskrivningen att missförstånd undviks. Exponeringen kan vara öppen eller tillhandahållas restriktivt allt efter omständigheter, vilja och förmåga.

Informationen ifråga, även om den kallas metadata, är ingenting annat än information som de potentiellt samverkande parterna är intresserade av att utbyta. Likheter med den typ av information vi talat om tidigare är slående. Alltså bör samma förfaringssätt gälla metadata. Metadata måste struktureras i enlighet med välformulerade, överenskomna begreppsmodeller. Gärna uttryckt i något standardiserat begreppsmodelleringspråk. Kunskap om vilket modelleringspråk som använts, samt tillgängligt operativt gränssnitt för att söka bland metadata måste också finnas, bland en hel del annat. Denna senare typ av information kan ses som ett slags metametadata (ursäkta terminologin), som förstås i sin tur behöver finnas tillgänglig – helst direkt på webben för att på så vis erbjuda automatik i sök- och hanteringsprocesser. Fram växer en semantisk webbmiljö. [18] [19]

Lägg till denna semantiska webbmiljö möjlighet att etablera roller som informationsmäklare och distributörer. Varför inte också avtalsexperter, analytiker, rådgivare, övervakare, garantier, inspektörer, skiljemän och en mängd andra roller som har funnit sin plats i vårt vanliga samhälle? Komplettera med möjlighet för dessa roller att skapas, ha ett liv och sedan avvecklas och fram träder något som skulle kunna liknas vid en dynamisk webbmiljö eller virtuellt samhälle. Visst låter det spännande.

E-Society II har i planerna för 2004/2005 avsatt tid för vidareutveckling av idéer kring en dynamisk webbmiljö. [20]

Rapportlista

Rapporter med fet stil finns upplagda på såväl e-Society IIs webbplats (www.skriver.nu/esociety) under "rapporter" och Serviams webbplats (www.serviam.se) under "stödfunktioner".

- 1 "Web Services – fakta och synpunkter", PowerPoint-bilder, 2003.
- 2 "Termen 'tjänst' – några perspektiv", 2004
- 3 "Generella verksamhetstjänster – finns dom?", 2004
- 4 "Att utveckla tjänster – lättare sagt än gjort?", 2004
- 5 "XML Schema – en översikt", 2004
- 6 "Begräppsmoelldriven informationssamverkan – databasansats", 2004
- 7 "Geografisk information - Beskrivningsspråk för begräppsmoellering", 2001
- 8 "The Binary Modeling Language", 2002
- 9 "Synpunkter på moellering av begräpp och termer", 2004
- 10 "Vokabulärsamordning – diskussion kring hantering av digitala lärresurser", 2004
- 11 "Conceptual modeling – some thoughts", 2004
- 12 "Begräppsmoelldriven informationssamverkan – utbytesansats", 2004
- 13 "Begräppsmoelldriven informationssamverkan – exponeringsansats", 2004
- 14 "Metadata – Vad, När, Hur, Varför?", 2003
- 15 "Metadata – What, When, How, Why?", 2003
- 16 "Mycket 'meta' är det", 2003
- 17 "A Conceptual Model Driven Semantic Web – a Binary Approach". Gemensam rubrik för ett antal rapporter, 2002
- 18 "Den semantiska webben – en introduktion"
- 19 "Konferensrapport från '1st European Semantic Web Symposium, May 2004'"
- 20 "A Virtual Digital Society – Beyond the Semantic Web". Gemensam rubrik för ett antal rapporter.