

INFORMATIONS- OCH DATAKVALITET

Staffan Ögren

Spridningsförbehåll:

Denna rapport får endast spridas och användas inom de organisationer som deltar som parter i TRIAD-projektet. ©TRIAD juni 1994

Innehåll

| | | |
|----------|--|-----------|
| 1 | Inledning och läsanvisning | 1 |
| 1.1 | Tack | 2 |
| 1.2 | Sammanfattning | 2 |
| 1.3 | Vad är data- och informationskvalitet? | 6 |
| 2 | Hur ska området informationskvalitet angripas? | 10 |
| 2.1 | Att ta hand om problemen allteftersom de kommer – "troubleshooting" | 11 |
| 2.2 | Medveten kravställning | 12 |
| 2.3 | Enklare åtgärder som påbyggnad till "troubleshooting" | 14 |
| 2.4 | Kalkylresonemang för val av angreppssätt | 14 |
| 2.4 | Fasta regler och normer | 16 |
| 3 | Väv ihop generellt kvalitetsarbete och arbete med informationskvalitet | 23 |
| 3.1 | Teman inom kvalitetsarbete som kan vara intressanta i samband med informationskvalitet | 25 |
| 3.2 | Praktisk vägledning inom kvalitetsområdet | 28 |
| 4 | Resultat med enkla åtgärder | 29 |
| 5 | Ställa krav på informationskvaliteten | 31 |
| 6 | Arbeta med informationsprocesser, ansvar och intressen | 53 |
| 6.1 | Långa och sammanflätade datakedjor ger kvalitetsproblem | 53 |
| 6.2 | 1:1:1:1-principen | 55 |
| 6.3 | Svårigheter att förena intressen | 56 |
| 6.4 | Kvalitetsstyrning | 64 |
| 7 | Mät och informera om informationskvaliteten | 67 |
| 7.1 | Mätning av informationskvalitet | 67 |
| 7.2 | Deklarationens grundmaterial | 71 |
| | Epilog | 76 |
| | Några ordförklaringar | 77 |
| | Referenser | 78 |

1 Inledning och läsanvisning

Vilken är den faktor eller det synsätt som kommer att få störst betydelse för utvecklingen av informationssystem under de närmaste fem åren?

Informations- och datakvalitet är en kandidat. Det är ett ämne som genomsyrar hela området informationssystem och dess tillämpningar och kan därför vara lite svårt att få grepp om.

När man talar om data- eller informationskvalitet är det olika slags problem eller aspekter på informationssystem som tas upp vid olika tillfällen. Begreppet kvalitet ses än som egenskaper hos data – rätt eller fel – än som frågan om hur principer för kvalitetsstyrning kan tillämpas på datasystem.

I denna rapport skärskådar vi i god IA¹-anda några olika aspekter på data- och informationskvalitet. Vi tar upp några centrala teman och ser efter hur man kan agera.

Rapporten innehåller olika uppslag till hur ett område kan angripas. Det ena angreppssättet behöver inte utesluta det andra. Men förslaget till läsaren är: välj *ett* av angreppssätten. till att börja med Belöningen hägrar i form av minskade kostnader för informationssystem och högre antal produktivitets- och kvalitetsmål som uppfylls i verksamheten. Kort sagt så minskar kostnaderna för kvalitetsbrister och konkurrenskraften blir bättre.

När man börjar arbeta är det inte nödvändigt att det som angrips verkligen är en informationskvalitetsfråga. Det viktiga är att frågan är central för verksamheten. Aspekten informationskvalitet är, som den läggs fram här, avsedd att vara till hjälp för att hitta en handlingsväg. Arbetet med informationskvalitet kommer att vara integrerat i arbete med DA och IA, systemutveckling och arbete med datasäkerhet.

I denna skrift begränsas ämnet till att i första hand behandla kvalitetsaspekter på traditionell databehandling, dvs tecken och sifferdata i strukturerad form. Framväxten av textmultimedia kommer att vidga ämnet kraftigt.

Kapitlen 1 - 4 ger en introduktion till ämnet och bör kunna läsas även av den som inte har djupare kunskaper inom databehandlingsområdet. Kapitel 5 - 7 ger en fördjupad behandling av några centrala områden.

¹ IA – informationsadministration. Den läsare som tidigare inte läst någon TRIAD-rapport och inte kommit i kontakt med förkortningarna DA och IA hänvisas till avsnittet *Några ordförklaringar* i slutet av skriften.

Framställningen varvas med skildringar av en tänkt verklighet. Avsikten är att frågorna ska illustreras på ett levande sätt.

1.1 Tack

En viktig och avgörande förutsättning för tillkomsten av denna rapport har varit resultaten från Peter Nilssons arbete inom området. Jag vill tacka honom för allt material han ställt till förfogande och även för att han tålmodigt har svarat på frågor och lämnat värdefulla synpunkter på utkast till skriften.

Sören Lindh förtjänar också ett särskilt tack att han oförtröttligt givit nya förslag till förbättringar i framställningen.

Andra som bidragit på olika sätt i arbetets olika skeden och som jag vill tacka är: Bertil Andersson, Bengt Carnö, Mauri Kemppi, Clary Lundahl, Karin Lundgren, Björn Thornadtsen, Göran Ledell, Aleksander Szyber och Ulf Åsén.

1.2 Sammanfattning

Data- och informationskvalitet kommer att beröra oss alltmer i och med informationssystemens allt större betydelse i vårt arbetsliv. Dessutom ökar betydelsen också i och med att arbetslivet i sig förändras och att möjligheterna till överföring av data mellan olika miljöer blir allt större. Vi går mot en elektronisk värld där direktkontakt med verkligheten ersätts av data från många olika och varierande källor, när och fjärran. Datas egenskaper – kvaliteten – måste uppfylla de krav som följer av en verksamhets behov och dessutom vara känd för alla berörda.

Huvudfrågan när det gäller data- och informationskvalitet är om data är lämpliga för det syfte vi använder dem till. Data kan aldrig ses som korrekta i sig, utan endast i förhållande till ett visst syfte. När nya data ska definieras måste syftet klarläggas entydigt och kraven tas fram utifrån detta syfte. För att kunna bedöma om befintliga data är lämpliga för ett visst syfte måste vi känna till syftet med den ursprungliga insamlingen, hur det uttrycks i olika modeller, definitioner, krav osv. Det är också nödvändigt att känna till i vilken utsträckning de data som är tillgängliga uppfyller kraven. Har vi tillgång till dessa data om data finns förutsättningar för användning och återanvändning. Det räcker alltså inte med att bara känna till t ex begreppsdefinitioner, även om dessa är en viktig del.

Ämnet kan hanteras på olika sätt. I en situation som är statisk och där vi får återkoppling på vår dataanvändning *kan* det vara tillräckligt att ta hand om kvalitetsproblem alltefter som de kommer. I många fall är det dock befogat med ett medvetet och planerat kvalitetsarbete med krav, åtgärder och uppföljning. Det gäller t ex i situationer när nya krav ställs, när konsekvenserna av fel i data eller missuppfattningar av data är allvarliga, eller när vi har en ambition att höja oss över mängden och uppnå "total kvalitet".

Informationen och dess kvalitet bör självklart ingå som en del av övriga kvalitetssatsningar. Från kvalitetsområdet finns det också en hel del uppslag att hämta till hur området data- och informationskvalitet ska behandlas, t ex utifrån begrepp som "Total Quality Management", "Kundorientering", "Gör rätt från början", "Kaizen" osv.

För att komma igång med arbetet kan en enkel ansats tillämpas. Den kan t ex innehålla användarforum eller ett kortare grupparbete bland personalen. Arbetet kan både ge uppslag till direkta förbättringar och fortsatt arbete.

För ett mer kvalificerat arbete med data- och informationskvalitet krävs en medveten kravställning som fast grund. Dessa krav måste i sin tur vila på fastställda krav för verksamheten i övrigt. Terminologin är ett viktigt verktyg och används för att uttrycka krav på kvalitet i noggrannhet, aktualitet, täckning, objektiva fel (feltolerans) osv.

Den kvalitet som data och information får uppkommer huvudsakligen på två sätt, dels genom den avbildningsmodell (datamodell) man har valt, dels genom det system som uppdaterar data utifrån förändringar i det som avbildas. Modellen och tillhörande definitioner, beskrivningar m m specificerar den *avsedda kvaliteten*, medan det arbetande systemet med dess instanser av data står för den *faktiska kvaliteten*. Båda aspekterna behöver beaktas i ett kvalitetsarbete.

Kraven på datakvalitet och informationskvalitet grundar sig på verksamhetens kvalitetskrav och produktivitetsambitioner. När kraven ska fastställas, värderas olika faktorer: behov hos primär- och sekundär användare, ekonomi, kravens stabilitet, risk- och effektanalys m m. I kravarbetet ingår även att planera för uppföljning av kvaliteten.

Genom långa dataflöden som flyter ihop och delar sig, sträcker sig över organisatoriska gränser osv uppkommer komplexa system som kan vara svåra att överblicka och kontrollera. Dessa flöden, som också kan kallas informationsprocesser, behöver kartläggas i alla led från ax till limpa. Olika parter ansvar och intressen behöver redas ut och kartläggas. Avtal och överenskommelser kan behöva upprättas mellan parterna. Återkopplingar behöver skapas i processerna för att kvaliteten ska kunna styras.

Allt kvalitetsarbete kräver mätning, så också data- och informationskvalitet. Mätning och beskrivning av data- och informationskvalitet kan göras på olika sätt: genom iakttagelse av registeregenskaper, genom att ett urval ur registret kontrolleras mot en säkrare källa eller genom att informationsbildningen kontrolleras, dvs att man undersöker om användarna tolkar data så som man avsett.

Dessa mätningar bör även kopplas till effekter av kvalitetsnivåer genom att man gör en koppling till mätningar av kvaliteten i själva verksamheten, dess mål och resultat.

Informationsanvändning med kvalitet kräver också att användarna får information om kvaliteten i tillgängliga data. Sådan information kan delges i kvalitetsdeklarationer. Deklarationen innehåller i sin mest fullständiga form en fullständig redovisning av avsedd och faktisk kvalitet men materialet måste sedan anpassas för olika mottagare, både till innehåll och distributionsform.

Ödestimmen slår

Du sitter framför PC:n en sen kväll och ska göra en kalkyl över en investering i ett nytt projekt. Om projektet lyckas är även *din* lycka gjord, men om det inte går bra kan du packa ihop och se dig om efter ett annat arbete. Fläkten i PC:n surrar hemtrevligt och arbetet går bra. Det finns MYCKET data som du med moderna verktyg kan hämta in från olika håll till ditt kalkylprogram och bearbeta kors och tvärs. Försäljningsstatistik, prisprognoser, marknadsdata, *allt* finns på fingertoppsavstånd. Fantastiskt! Du behöver inte byta ett enda ord med dem som har samlat in alla dessa data för att få en fullständig bild av världen.

Kalkylresultatet ser fint ut. Du provar lite känslighetsanalyser. Ja, inte så farligt, vi klarar en uppgång i kronkursen med 15%.

Nu kommer Svenne in i rummet. Jag *har* inte tid att prata med honom, tänker du.

”Hej du, kvar så här sent? Ja det ju bra för karriären. Du har försäljningsstatistiken för korsfabulatorer uppe ser jag.

Men vet du att kvaliteten är lite dålig på dom siffrorna? Dom flesta uppgifter som lämnas in är nog ganska överdrivna, marknaden är inte så där stor för såna grejer. Men man vill ju att det ska se bra ut. Och sådär 30% av dom där fabulatorerna som säljs är ju egentligen intuitionsdatabaser, men klassas som fabulatorer ändå. Den sorten fanns ju inte när klassificeringsmallen gjordes. Men det är väl inte avgörande för dig hoppas jag. Dom som använder de där siffrorna är mest intresserade av totalen på ett ungefär, och det framgår ju. Jaha, dags att gå hem, hörnu, sitt inte kvar för länge och räkna nu. Hej då!”

Du sjunker nedåt i stolen. Om 30% av produkterna i statistiken är av den nya sorten som egentligen borde ha en egen klassifikation. Jag måste räkna om, då blir det... Vänta nu, hur kan jag veta att det verkligen är 30% ?

Och resten av siffrorna från den statistiska databasen då, hur är det med dom?

Golvet känns mjukt under fötterna. Siffrorna börjar dansa runt på skärmen.

Ja, så kanske det kan se ut, en sen kväll på kontoret.

Låt oss gå tillbaka och fundera över vad situationen innebär.

Ditt problem i exemplet är till en början att du inte är medveten om svagheterna i informationskvaliteten eller brister i data. När du med Svennes benägna hjälp blir medveten om detta har du plötsligt fått ett annat problem. Du vet eller misstänker att det finns ett fel, men vet inte hur stort det är. Dessutom verkar det som siffrorna är framtagna med ett delvis annat syfte än ditt och då är det tveksamt om siffrorna mäter rätt sak.

Problemsituationen illustrerar två aspekter på kvaliteten: om data är korrekta i sig och om data, under förutsättning att de är korrekta, passar för det aktuella syftet.

Terminologiruta

”Validitet beskriver om lagrade data är korrekta utifrån vad som gäller för valda objekt.”

”Relevans beskriver om lagrade data är ändamålsenliga, passar till det problem man vill lösa.”

(ITS 94)

Du har riskerat att falla offer för myten om att alla data är sanna. Egentligen vet du förstås bättre. Säkert har du redan erfarenhet av opålitliga data. Men sådant är lätt att glömma. Särskilt om erfarenheterna ligger en bit tillbaka i tiden och om siffrorna på skärmen stämmer med det du förväntar dig att se eller det du *önskar* se.

Du har blivit osäker på om du kan lita på data, men har inget mått på hur pass opålitliga data är. Visste du det så skulle du kunna göra en ny känslighetsanalys.

Du börjar undra över vad data- och informationskvalitet egentligen är.

1.3 Vad är data- och informationskvalitet?

Några aspekter på frågan om vad datakvalitet och informationskvalitet är behandlas här. Frågan behandlas mer utförligt i avsnittet *Ställ krav på informationskvaliteten*.

Syftet är alltid grunden

När data samlas in görs detta i ett speciellt syfte. Att ha syftet klarlagt är grundbulten för allt arbete med data- och informationskvalitet. Syftet kan vara mer eller mindre specifikt uttänkt och presenterat men det finns alltid där. Om syftet är allmänt och ospecificerat så finns risken att de data som samlas in inte blir lämpliga för någon användning alls.

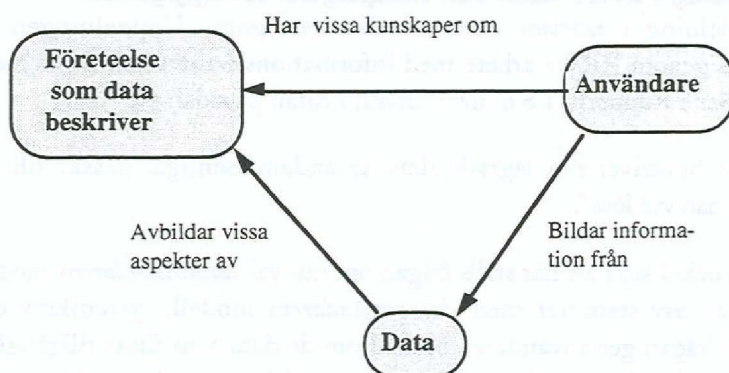
Data som är mycket specifikt anpassade till ett syfte kommer inte att passa alls i ett annat syfte. Ett exempel på detta är hur data med namnet "sjukpenninggrundande inkomst" använts i felaktigt syfte. Detta beskrivs i ett senare kapitel.

Om data är grundläggande basfakta som är baserade på allmänt kända och accepterade definitioner, kan de vara användbara i flera syften. Syftet är i det fallet just att skaffa sig ett grundläggande lager av data som via aggregeringar eller annan behandling kan användas i flera syften. Begränsningarna för i vilka syften de kan användas måste dock vara kända.

I själva insamlingen av data har också syftet en betydelse. Den som lämnar uppgifter kommer, i sitt uppgiftslämnande, att påverkas av sin uppfattning av syftet med insamlandet. Den som utför själva insamlandet av data på fältet påverkas av sin uppfattning av syftet i sin bedömning av hur viktigt det t ex är med mycket noggranna mätningar. Dessa aspekter måste beaktas även när man avser att bygga upp ett datalager med basfakta.

Syftet måste alltså vara känt för alla som kommer i kontakt med data, från det att de samlas in, till dess att de används.

Användaren, data och information



Figur 1

Användaren har redan vissa grundkunskaper om det som data beskriver, eller avbildar. Användaren utgår från *sitt* syfte och sina tidigare kunskaper, iakttar företeelsen indirekt via data och bildar information. Detta sker naturligtvis på gott och ont. Datasystemet kan överbrygga tid och rum och stödja kommunikation, men är begränsat i fråga om vilka aspekter som kan iakttas eller kommuniceras.

Fel i register

Det som man kommer att tänka på först i samband med datakvalitet är vanligen frågan om hur många data i ett visst register som är korrekta respektive felaktiga.

I vissa fall är det ganska självklart vad som är rätt och fel. Det kan t ex handla om uppenbara och odiskutabla fel som att siffror i ett telefonnummer har blivit omkastade, eller att ett mätvärde skiljer sig 200% från det rätta. I andra fall är gränsen mindre tydlig. Det kan handla om fel som ligger i närheten av någon gräns och där en bedömning mot en norm måste göras, t ex på vilken nivå i ett klassificeringsschema som en klassificering måste vara rätt². Syftet med dataanvändningen är återigen en avgörande faktor.

² Exempelvis företag klassificeras i vissa sammanhang enligt "Svenskt Näringslivsindex", den s k SNI-koden. Koden är femställig och framtagen för statistiska ändamål. Detaljeringsgraden fem siffror kan vara onödigt stor för en del ändamål där det kanske räcker med treställig kod. I den treställiga koden kan man t ex få reda på om ett företag sysslar med med samfärdsel i form av "Land-, sjö- och lufttransport" eller "Transportförmedling, lagringstjänster". Med en femställig kod kan man i gruppen "Land-, sjö- och lufttransport" skilja mellan järnvägar, spårvägar, bussrörelse, taxirörelse, lastbilsäkerier, transport i rörsystem osv.

Uppdelning i korrekthet och lämplighet för uppgiften

En uppdelning i *relevans* och *validitet* kan göras. Uppdelningen har utvecklats genom RRV:s arbete med informationskvalitet och finns med i definitionen i Rapport ITS 6 (även citerat i rutan på sidan 6):

”Relevans beskriver om lagrade data är ändamålsenliga, passar till det problem man vill lösa”.

Man kan också säga att här ställs frågan om hur väl *datainsamlarens* modell, systemkrav osv stämmer med *dataanvändarens* modell, systemkrav osv. Svaret på frågan ger användaren besked om de data som finns tillgängliga är lämpliga för den avsedda användningen. Eller uttryckt med en enkel fråga: ”Beskriver dessa data rätt sak på det sätt som jag vill?”

”Validitet beskriver om lagrade data är korrekta utifrån vad som gäller för valda objekt.”.

Man kan säga att det handlar om huruvida data är korrekta med utgångspunkt i insamlarens modell med dess syften, definitioner och toleranser.

Uppdelningen är värdefull eftersom den pekar på två grundläggande perspektiv. I det ena perspektivet betraktas datas korrekthet i sig, i det andra lämpligheten hos data för det specifika syfte en viss användare har med sin användning. När kvaliteten diskuteras bör man försöka fokusera perspektiven vart och ett för sig efter varandra.

Efter detta görs en sammanvägning eftersom uppdelningen inte är uteslutande. Korrektheten (validiteten) hos data påverkar t ex dess lämplighet för en viss användning (relevansen). Det är alltså möjligt att tänka sig en situation där det råder oklarhet om en brist i kvaliteten i första hand ska hänföras till validitet (t ex mätfel) eller relevans (t ex avvikande modell eller syfte). Att avgöra vilket är kanske inte nödvändigt, det viktiga är att uppmärksamma båda aspekterna innan en slutsats dras.

Ordvalet validitet och relevans kan uppfattas som abstrakt. I dagligt tal kan man använda enklare ord som t ex korrekthet i stället för validitet. Även om olika termer används så bör man dock inte tappa bort skillnaden i perspektiv mellan de två begreppen.

Data i sig och data för mig

Angränsande till frågan om syfte kan man dela upp kvalitetsbegreppet i två delar, en objektiv och en subjektiv del. Den objektiva delen är egenskaper hos data i sig och dess förhållande till det som data handlar om. Det handlar om egenskaper som går att beskriva på ett faktamässigt sätt genom specifikationer, mätningar osv, och som inte innebär en värdering i förhållande till ett visst syfte. Sådana fakta kan vara vilken mätskala som har använts, mättidpunkt osv.

Den subjektiva eller relativa delen är den bedömning som en användare gör av informationens användbarhet "för mig" eller lämplighet för ett visst syfte. En sådan bedömning kan t ex uttryckas som att "data är för gamla", "fel aspekter belyses", eller "oacceptabelt många fel i kundregistret".

Den objektiva beskrivningen av egenskaper hos data eller register, i form av faktauppgifter, är grunden för användarens bedömning av om data är lämpliga för ett visst syfte.

Man brukar beskriva data- eller informationskvalitet med termerna noggrannhet, aktualitet, täckning, objektiva fel osv. Detta och ytterligare aspekter på vad data eller informationskvalitet är behandlas vidare i kapitel 5, *Ställa krav på informationskvaliteten*.

2 Hur ska området informationskvalitet angripas?

”Nu är det en kund som ringer igen och skäller på stackars Lisa för en faktura som han fått *efter* det att halva betalningstiden har gått ut” sade Lasse. ”Adressen var förstås fel *trots* att de meddelat sin nya adress i *god tid* före flyttningen. Det är verkligen hög tid att vi gör som Elisabeth föreslog för ett halvår sedan, lägger upp ett kvalitetssystem för data och information. Vi förlorar kunder på den här oredan”.

Systemchefen blev först röd i ansiktet, men sade sedan med tillkämpat lugn:

”Jomen förstår du. Det här med framtida adressändringar, det är ett slags undantagshantering som måste skötas manuellt. Om vi lägger in den nya adressen med en gång så skickas fakturorna dit *innan* kunden har flyttat, och det är väl inte så bra, eller hur?”

Systemchefen började nu se lite mer säker ut. ”Och har du den minsta aning om vad ett kvalitetssystem kostar? Dyra konsulter, certifiering, kvalitetsrevisorer som ska snoka här flera gånger om året. Formaliteter för sin egen skull för att lösa en massa problem som inte finns. Allt för att en enda kund fått sin faktura sent. Ta det du lugnt. Det där kan säkert Svenne fixa med en enkel ändring i programmet. Vi tar tag i problemen när de uppkommer, det är ett utmärkt och billigt kvalitetssystem. Och jag åtar mig dessutom att *själv* ringa upp den kund som har fått fakturan sent och förklara situationen. Vi kan skicka en flott bukett blommor och en chokladask också så kommer dom att *älska* oss i stället för att vara sura och vi vänder felet till vår fördel i stället. God kundvård kallar jag det!”

Ja, vem har rätt? Är det Lasse och Elisabeth, eller systemchefen? Kan Svenne fixa allt till det bästa? Räcker det med att lösa problem allt eftersom man upptäcker dem, eller behövs det en medveten satsning? Svaret är kanske inte helt självklart. Kanske finns det också dolda kvalitetsbristkostnader som bara kan upptäckas genom en medveten genomgång?

Antag att det finns ett begynnande intresse för data- eller informationskvalitet. Frågan blir då "Hur ska vi angripa det?" Möjligen har man också börjat fråga "Hanterar vi redan dessa frågor på ett rimligt bra sätt?" "Fungerar det tillräckligt bra redan?" "Vilka är felkällorna?"

Frågan är nu om man ska fortsätta med att ta hand om problemen allt eftersom de upptäcks, eller om man ska tillämpa en medveten kravställning.

2.1 Att ta hand om problemen allteftersom de kommer – "troubleshooting"

Det vanligaste inom en löpande verksamhet är troligen att man åtgärdar problem när de uppkommer och inte ställer upp specifika krav på kvaliteten uttryckt i ett antal centrala variabler. Detta kanske inte är någon medveten strategi för att hantera data- eller informationskvalitet. Det betraktas som en del av det normala, dagliga arbetet med problemlösning och man sätter inte någon särskild etikett på det. Det kan sägas motsvara mycket av den inställning som var rådande även inom övriga områden innan kvalitetsfrågorna började behandlas på ett mer ordnat och nytt sätt.

När kan troubleshooting vara tillfyllest?

För att besvara frågan om när troubleshooting är tillfyllest i dataanvändning kan definitionen från Rapport ITS 6 bilda utgångspunkt:

Terminologiruta

"**Informationskvalitet**, användbarhet, också med hänsyn till dess effekt, av information för en given användare och ett givet problem"³.

ITS 94

³ Kommentar: Problem bör kunna uppfattas som synonymt med syfte eller ändamål; vad det är man vill åstadkomma eller uppnå med sin användning av informationen.

Angreppssättet att ta hand om felet när de uppkommer, kan vara tillfyllest om

- användningen är regelbunden och förutsägbar, dvs om vi vet vem som är användaren och vad som är syftet (problemet) och detta inte ändras
- effekterna av användningen är kända dvs ger synbara effekter som uppmärksammas. Det finns en återkoppling av eventuella problem på grund av fel)
- de fel och effekter av fel som uppstår inte är av en allvarlig natur

Om vi t ex använder ett adressregister för utskick av inbjudningar till en konferens kan det hända att vi får ett litet antal returerna som beror på inaktuella adresser. Returerna ger en återkoppling på en aspekt av hur pass användbart registret är, nämligen om breven överhuvudtaget kommer fram till adressaten. Mängden returerna är inte avskräckande men en del av dem kommer efter rätt så lång tid. Totalt kommer det in så pass många anmälningar till vår konferens så vi kan anta att registret har en acceptabel kvalitet så tillvida att postadresserna är korrekta och att rätt personer nås. Vi brukar skicka ut inbjudningar då och då, alltså en regelbunden verksamhet. Vi har hyfsad kontroll på användbarheten genom returerna och effekten av felet är inte så oroande eller allvariga.

Detta kan sägas vara en situation som kännetecknas av att vara statisk i fråga om informationsanvändandet, och där riskerna för fel som man inte har full kontroll över är små. De stora fel som skulle kunna inträffa har antagligen redan inträffat.

2.2 Medveten kravställning

Det mer kvalificerade – och därmed resurskrävande – alternativet är att medvetet ta ställning till kravnivåer för olika variabler, t ex i fråga om noggrannhet beträffande mätvärden eller hur många objektiva fel av ett visst slag som är acceptabla. Detta ställningstagande kan i sin tur kräva utredningsinsatser för att avgöra hur informationskvaliteten påverkar kvalitet och produktivitet i själva verksamheten. Det kräver insikt i ämnets detaljer och tillgång till en begreppsapparat som ett verktyg i arbetet.

RRV tar i "Rätt data?" (RRV 90) upp bristen på kravställning som ett huvudproblem: "Användaren ställer inte tillräckligt preciserade krav, baserade på en riskanalys. Därför finns det risk att data av undermålig kvalitet används som beslutsunderlag". RRV har också i samma skrift konstaterat att "Det finns fall där användaren inte kände till att data av undermålig kvalitet användes."

När bör en medveten kravställning tillämpas?

Antag i stället att vi vill börja använda uppgifter från det tidigare nämnda adressregistret i ett nytt syfte, nämligen att sända ut fakturor och betalningspåminnelser. Innan vi börjar med det så försöker vi dock göra en uppskattning av hur många postadresser som har varit fel. Vi gör en gissning på att fem procent av adresserna är så pass felaktiga att brevet kommer i retur, i en del fall efter rätt så lång tid⁴, vad händer då?

Risken för allvarliga konsekvenser är stor. Det finns en risk att inkassoföretag kopplas in på kunder som inte har fått sin faktura ännu.

När det gäller ett sådant syfte drar vi slutsatsen att vi måste ha ett eget, välkontrollerat register, där t ex adressen i registret kontrolleras mot den adress kunden anger på beställning eller konferensanmälan.

Detta är då ett exempel på en situation där en mer medveten kravställning är motiverad.

Situationer när kraven på informationskvalitet behöver behandlas på ett medvetet sätt är när:

- nya krav lyfts upp från omvärlden, från kunder eller andra, eller internt t ex från ekonomiadministrationen
- verksamheten där data ingår och används ändras
- nya system ska byggas
- större förändringar ska göras i befintliga system
- IA eller DA ska introduceras i organisationen
- data används för nya syften
- det finns höjda ambitioner i verksamheten i fråga om kvalitet och produktivitet
- man är beroende av en eller flera organisationer för att få information

Vi kan också finna karakteristiska drag i situationen genom att vända på listan ovan:

- Användningen är inte regelbunden och förutsägbar, dvs data är generellt tillgängliga för olika syften. Vi vet inte säkert vem som är användare av vissa data, det varierar. Vilka datamängder som används varierar också. Det kan vara data med internt eller externt ursprung. Syftet med användningen kan vara olika från gång till gång.

⁴ När det tar tid så har adressundersökningen hos Posten gjort ett försök först. I en hel del fall leder denna undersökning till att försändelser med en inte helt korrekt adress kommer fram, men de blir försenade.

- Effekterna av användningen är inte kända, dvs det saknas återkoppling av eventuella problem på grund av fel. Effekterna är inte så påtagliga och starka att de uppmärksammas med krav på åtgärder.
- De fel och effekter av fel som uppstår kan vara av allvarlig natur. Eftersom användningen av data inte är regelbunden kan helt nya problem uppstå vid helt nya användningar.

I inledningen beskrevs en situation med stor risk. Data skulle användas för att man skulle kunna bedöma en ny satsning. Eftersom projektet var nytt fanns inte någon möjlighet till återkoppling från just den användningen. En bedömning av informationskvalitetens betydelse behöver därför göras i förväg.

Hälsan tiger heller inte alltid still. När saker och ting förefaller att fungera bra innebär det inte att man kan slå sig till ro. Det kan vara dags att ställa frågan om de kan fungera ännu bättre; och om informationskvalitet kan vara en faktor att ta med i diskussionen. Även i den reguljära situationen kan det finnas kvalitetsproblem som man vant sig vid och som det kan löna sig att ta en titt på.

2.3 Enklare åtgärder som påbyggnad till "troubleshooting"

Ett alternativ till "troubleshooting" är att genomföra några enklare åtgärder till begränsad kostnad och ansträngning. En sådan satsning kan ge några konkreta resultat och även mer fakta i form av svar eller indikationer på hur det står till med data- och informationskvaliteten. Det kan t ex finnas problem som ännu inte ger tydliga symptom. Dessa fakta kan vara till hjälp när man beslutar om att hålla kvar den tidigare ambitionsnivån eller ändra den. Några sådana enklare åtgärder beskrivs i nästa kapitel.

2.4 Kalkylresonemang för val av angreppssätt

Ett resonemang om pengar kan ge ledning i valet mellan att ta problemen som de kommer och att ställa krav på ett mer planerat och medvetet sätt. Penningfrågorna handlar i grunden om två saker: *Vad är det som kostar?* och *Vad är det som ger intäkter?* Mer konkret kan de t ex formuleras så här:

- Är det förenat med stora kostnader att samla in och behandla den aktuella informationen?

Då kan det vara en god idé att undersöka om det går att klara sig med en mindre kostsam kvalitetsnivå. Man kan minska onödiga kostnader för kvalitetsnivåer som inte efterfrågas och inte leder till något mervärde i senare led.

- Har man gjort en uppskattning av vad felen i verksamheten kostar, direkt och indirekt, eller finns det en intuitiv uppfattning om att felkostnaderna är betydande?

Vad är det t ex som kostar:

- *direkt* – i form av minskad effektivitet och produktivitet i verksamheten, kostnader för felrättning, resurser som förbrukas på att göra fel, resurser som inte utnyttjas optimalt
- *indirekt* – i form av missnöjda kunder som inte hör av sig utan bara försvinner, kostnader för att ge fel erbjudande till kunden, avtalsbrott, lagbrott osv

Från en sådan uppskattning kan man gå tillbaka och se efter vilken roll informationen spelar. Hur är påverkan om en medveten styrning av data- och informationskvalitet helt saknas? Hur påverkas kostnaderna ovan av att specificerade kvalitetsnivåer inte upprätthålls? Om uppskattningen inte är gjord kanske den bör göras.

Kalkylexempel

Inom Televerket (som det då hette) genomfördes mellan år 1986 och 1990 ett datakvalitetsprojekt med titeln "Datakvalitet i själva verket". Man arbetade med de system som dokumenterar näten, särskilt lokalnät, och med planeringsverktyg.

Inom området lokalnät ledde projektet till årliga kostnadsminskningar på minst 35 Mkr. Projektet gav dessutom en engångsintäkt på cirka 90 Mkr, genom att man hittade ledig kapacitet i nätet som var registrerad som upptagen. Kostnaderna för projektet var totalt 60 Mkr fram till 1991 (Televerket 91).

Av Postens kostnader för fel i adresser finns det en risk för att 5 - 10 miljoner kronor per år ⁵ beror på fel i det statliga systemet för adressinformation ⁶ (RRV 92).

Direkta merkostnader för att hantera konsekvenser av felaktiga adresser hos *en* myndighet, CSN uppgår till 200 000 kr per år. De totala kostnaderna är dock sannolikt större. Exempelvis kan läntagarna drabbas av förseningsavgifter (RRV 92).

⁵ RRV betonar osäkerheten i siffran på grund av litet urval.

⁶ Det är viktigt att notera att detta system, så som det beskrivs i RRV:s rapport, inte är ett system som har planerats, beslutats av en ansvarig, konstruerats och satts i drift. Systemet har bara vuxit fram och SAIS (Statens adressinformationssystem) är en konstruktion som RRV har gjort för att kunna göra en helhetsstudie av hur adressinformationen hanteras i staten.

- Finns det vinster att göra i verksamheten eller kan man få konkurrensfördelar med informationskvalitet, t ex genom att vara den som har den mest tillförlitliga informationen, den mest aktuella, den mest fullständiga, den mest förädlade osv?

Dessa vinstmöjligheter är inte alltid uppenbara. Om vinsten ska tas hem genom att man skaffar sig en helt ny konkurrensfördel, som idag inte utnyttjas alls av någon aktör, kan ett seriöst idéarbete behövas för att hitta möjligheterna.

- Finns det risker för stora kostnader om problem med informationskvaliteten skulle uppstå?

Detta är riskanalyser som ofta görs som en del av normalt datasäkerhetsarbete. Se efter om något sådant material redan finns och om alla aspekter på informationskvalitet är täckta.

Detta var exempel på några frågor. Givetvis finns det fler som kan och bör formuleras utifrån den specifika verksamhet som ska granskas.

Materialet från dessa kalkylresonemang kan återanvändas när man ska ta ställning till kravnivåer. Ytterligare avsnitt om ekonomi följer i avsnittet *Ställ krav på informationskvaliteten*.

2.4 Fasta regler och normer

Inom alla organisationer finns det fasta värden eller normer som måste hållas. Det kan röra sig om lagar och regler men även om de fasta övertygelser som finns i organisationskulturen som t ex att en kund eller klient alltid ska behandlas på ett visst sätt oavsett vem det är. Om brister i informationskvaliteten gör att dessa normer inte kan hållas, är det nödvändigt att ställa medvetna krav.

För myndigheter och andra offentliga organisationer kan man behöva undersöka effekterna på organisationens sätt att utföra sin uppgift vad gäller t ex brister i rättssäkerheten vid myndighetsutövning. I detta ingår dels den specifika uppgift som varje myndighet sköter, dels de regler som gäller för hur uppgifterna ska utföras. Till det senare hör t ex frågor om offentlighet och integritet. I princip måste man utgå från att myndigheten utför sin uppgift med beslutad kvalitet och upprätthållen rättssäkerhet oavsett vad det kostar.

Det finns alltså regler för hur myndigheter ska arbeta som förhindrar vissa tänkbara rationaliseringar och därmed besparingar. Ett problem kan vara att det inte går att få klara och entydiga besked om vilken kvalitet verksamheten ska hålla i form av t ex resultatmått och mätbara kvalitetsmål. Grunden för att fastställa krav på informationskvalitet blir då lös. Att åtgärda brister inom detta område kan innebära, men behöver inte innebära, ökade kostnader för myndigheten.

Kostnader kan också orsakas av brister i myndighetsutövningen, t ex i hanteringen av överklaganden, processföringen hos myndigheten osv. Förbättrad kvalitet på myndighetens uppgifter och på upprätthållandet av rättssäkerheten kan också vara ekonomiskt fördelaktigt, men i detta fall är det inte ekonomin som motiverar insatser för bättre kvalitet.

För företag skulle man kanske kunna tro att allt är kalkylerbart, men även här finns det säkert områden där man snarare sätter upp en viss norm eller regel som helt enkelt ska följas än att man gör en kalkyl. Det gäller att följa lagar som t ex datalagen, men även egna etiska kvalitetsnormer som man anser är nödvändiga för att företagets goda rykte ska upprätthållas. Det finns naturligtvis ekonomiska kopplingar till företagets överlevnad på lång sikt men det kan vara svårt att identifiera och kvantifiera alla de orsaks-samband som behövs för att t ex göra en kalkylerad avvägning mellan de kostnader och de vinster som förknippas med ett gott namn.

Tendenser som ger skäl att uppmärksamma informations- och datakvaliteten

För att ge ytterligare uppslag i bedömningen av om informationskvaliteten ska uppmärksammas ges här en genomgång av några aktuella tendenser som har en koppling till ämnet.

Förändringar och nya krav i organisationer

Stora och allt snabbare förändringar ger höjda krav i morgondagens organisation. IT-utveckling med datautbyte är både en pådrivande faktor och en möjliggörare.

Decentralisering är idag vanlig. Detta kräver mer datautbyte mellan organisationens olika delar för rapportering och styrning⁷. Lokala enheter behöver ha tillgång till centrala data och till varandras data. Centralt behöver man ha snabbare besked om prestationer och resursförbrukning lokalt för att kunna styra helheten dynamiskt. Decentralisering kan också leda till dubblering av system. Flera register som avser samma sak byggs upp. Detta innebär onödiga kostnader och risk för motstridiga uppgifter. Organisatoriska förändringar som decentralisering sker i en allt snabbare takt varför perioder med regelbunden och förutsägbar verksamhet blir allt kortare.

Antalet mellanchefer minskar. Dessa chefers arbete med att sila, bearbeta, styra och anpassa information ersätts av datasystem t ex av typen Executive Information Systems (EIS). Högre chefer och andra nyckelpersoner måste direkt kunna tolka data med varierande bakgrund.

Nya organisationer kräver att medarbetarna har överblick och bredd. Skärpta servicekrav – serviceflexibilitet – innebär att var och en måste kunna representera hela företaget eller organisationen som helhet. Det gäller att kunna hjälpa till med att formulera frågor åt kunderna och sedan skicka frågorna vidare till någon annan, sortera problem som man kan lösa på egen hand/referera till andra och ofta själv återkomma med besked. Detta i stället för att hantera endast en viss typ av ärenden. Då behövs mer bakgrundsinformation från olika delar av organisationen. Var och en som är ute på fältet eller har kundkontakter på annat sätt måste kunna komma åt olika slags data som inte hör till det egna specialområdet. Det rör sig då om data från många olika källor och som är skapade i olika sammanhang. Individerna måste kunna förstå vad dessa data står för och även kunna värdera deras lämplighet i övrigt.

Även specialisten måste vara flexibel, kunna hoppa in på nya arbetsuppgifter och yrkesroller och då använda och förstå de data som den rollen använder. I dagens läge tar man inte gärna in vikarier utan den ordinarie personalen får rycka in i nya roller när det behövs.

Krympande resurser och rationaliseringstryck i många verksamheter i förening med tekniska möjligheter ger större krav på att inte själv samla in information utan lita till andra källor.

⁷ I RRV 1991:11 talar man t ex om samverkande lokala ADB-system som utbyter data med system i andra myndigheter.

Den offentliga sektorn kommer i större utsträckning att få betala för den information den behöver skaffa utifrån. Det gäller då att olika verksamheter inom den offentliga sektorn blir goda beställare/kravställare även när det gäller informationskvalitet. Dessa verksamheter kommer därför att behöva mycket bra beskrivningar av den information de erbjuds att köpa för att kunna bedöma dess användbarhet och värde.

ADB förändras i det nya arbetslivet från att vara en separat industriell process till att bli en integrerad del av verksamheten. Under den tid de stora datorcentralerna dominerade var det särskild personal som kodade och matade in data i stora mängder. Detta gav möjlighet att organisera kontroller på ett systematiskt sätt. Nu är stordatorn en anonym server. Data matas direkt in i lokala system ärendeför ärende, som en del i handläggningen. Kontroller måste ske på andra sätt än vad som är möjligt vid mängdregistrering (t ex buntsummor). Ansvar för tillämpningssystem i sin helhet hamnar idag hos personal som inte alltid har någon djupare ADB-utbildning och som ska klara av att hantera det tekniska stödet vid sidan om det egna fackområdet.

"The virtual corporation" – den virtuella organisationen byggs upp kring informationsnät som stöd för allianser med andra organisationer. Inom en större organisation kan det handla om ett "intraprenörprojekt"⁸ där en eller några eldsjälar får bygga upp något nytt. Det nya är fritt från företagets standardmönster och ligger utanför den ordinarie organisationen. I den virtuella organisationen finns begränsade möjligheter att lita till tidigare erfarenheter av dataanvändning. Flexibiliteten måste vara stor, data-systemen får inte låsa in verksamheten i ett visst handlingsmönster. Och om datakvaliteten inte är under kontroll...

Målstyrning ställer krav på allt mer resultatriktad information. Ofta lanseras ADB-lösningar för ledningsinformation såsom EIS, Executive Information Systems. I sådana system sammanställs information från många olika håll. De speciella kvalitetskraven för dessa tillämpningar måste vara kända och det måste även kvaliteten vara hos den information som används. Kvalitetsavvägningar kommer att behöva göras för information som ska användas, såväl i det operativa arbetet som för styrning och uppföljning.

⁸ "Intraprenör" kan utläsas "intern entreprenör". Man försöker tillgodogöra sig entreprenöranda inom organisationen. Personer som annars skulle ha behövt söka sig utanför organisationen för att kunna förverkliga sina initiativ ges utrymme att agera internt.

Affärsframgång kräver aktörer. I utvecklingen som beskrivs ovan får användaren en ny roll. Det handlar inte längre om att vara "en kugge i maskineriet" och reagera när data anländer. Användaren blir en *aktör* som strävar efter att nå ett mål och som söker och bearbetar data efter egen arbetsstil och situationens krav. Aktören måste ha mandat, frihet i sitt agerande och säker kunskap om kvaliteten hos de data han eller hon använder⁹.

Återanvändning

Återanvändning av data, definitioner, modeller osv är något som aktualiseras alltmer och därigenom är det också en kommande tendens. Att göra allting från början i alla utvecklingsprojekt är inte ekonomiskt och skapar dessutom risker för att t ex data från olika källor inte blir jämförbara. Återanvändningen skapar dock även kvalitetsrisker t ex när data som samlats in i ett syfte ska användas i ett nytt syfte. Då krävs kontroller för att säkra kvaliteten och även utförliga beskrivningar – *deklarationer* – av det som ska återanvändas.

Flödet av data underlättas

Det blir allt lättare att föra över data eftersom tekniska hinder tas bort och utbudet ökar. De elektroniska motorvägarna asfalteras i rask takt. Överförda data, som kanske kommer från miljöer långt bort ifrån eller från vårt nya närområde EU, kan ha ett informationsinnehåll som är oklart och kvaliteten kan vara okänd.

Exempel på sådant som undanröjer hinder för att överföra data är pågående standardisering och öppna system, framför allt med tjänster inom data-kommunikation¹⁰. Med standardiserade produkter har utbytet av data med omvärlden blivit lättare än förut.

Detta i kombination med EES och EU-integrationen kommer att leda till att data hämtas från olika, externa källor i ett intensivare datautbyte mellan olika delar av organisationer och mellan organisationer. En del kommer att hämtas långt bort ifrån – från sammanhang som vi har begränsad kunskap om.

⁹ Temat "från användare till aktör" utvecklas i TRIAD-rapporterna V4, "Den gemensamma informationsmarknaden" och V6 "Det kundeffektiva företaget".

¹⁰ T ex Internet och Stattel-initiativet inom staten.

Det blir allt vanligare med verktyg och system som gör det möjligt att ganska lätt sammanställa information från många olika källor internt och externt. Exempel på detta ¹¹ beskrivs i andra TRIAD-rapporter.

Förbättrad kommunikationsteknik, standarder osv möjliggör att man slipper att dubbellagra information på många ställen, slipper att sprida kopior och i stället kan hämta in informationen precis i det ögonblick den behövs.

Att vara säker på *vad* man får fram på sin bildskärm med stöd av denna utveckling blir allt viktigare.

En elektronisk värld

Verkligheten förändras med en tendens som innebär att det är databasen som gäller. Vi får en allt större andel "administrativ verklighet" eller "abstrakt verklighet" som innebär att data inte kan kontrolleras mot en fysisk verklighet, precis som våra sedlar inte längre garanteras av någon guldreserv i riksbanken. Detta innebär större svårigheter att kontrollera och större risk för att fel inte upptäcks spontant. (Tänk om felet också har spridits till den databas man kontrollerar mot?)

Datavärlden blir alltmer illusorisk med grafik, multimedia, virtuell verklighet osv. Det gör den svårare att skilja från verkligheten eller den sanna bilden av verkligheten. Det jag har sett med egna ögon tror jag på. Men vad händer nu om det visar sig att det jag ser inte stämmer? Användarens, ibland felaktiga, uppfattning att alla data är korrekta kan stärkas av att den illustrerande bilden ser mycket illusorisk ut.

När data fjärras från sitt ursprungliga sammanhang blir datakvalitet och informationskvalitet allt viktigare

Så länge som datasystemet används inom en viss verksamhet sker kvalitetskontroller mer eller mindre automatiskt, och kanske utan att man är direkt medveten om det. Genom att man själv har direktkontakt med det som data beskriver får man en allmän uppfattning om datakvaliteten. Användarna har därmed en någorlunda klar bild av i vilken utsträckning man kan lita på och använda den information data ger. Genom dessa kunskaper hos användarna och genom improvisationer kommer verksamheten att fungera någorlunda bra trots vissa brister i data.

¹¹ Text Information warehouse, hybris-debris och infofråga.

När data skapats i och för en viss miljö och även används i samma miljö finns förutsättningar som kompenserar för fel och missuppfattningar. När användaren betraktar data utifrån sin miljökunskap och faktakännedom ringer varningsklockor i huvudet när något är fel eller tvetydigt. När data används i en annan miljö saknas dessa förutsättningar till stor del. En osäkerhet om kvaliteten uppstår. Denna osäkerhet kan leda antingen till att man godtar data som de är och litar på dem till hundra procent eller också till en alltför stor skepsis om man upptäcker några enstaka fel. Man måste *vet*a hurdan kvaliteten är.

Om och när data ska användas på annat håll än där de har skapats krävs betydligt mer "data om data"¹², för att data inte ska missuppfattas. Detta gäller inte bara begreppsdefinitioner och dylikt utan även uppgifter om ursprunglig källa, hur data har behandlats på vägen mellan källan och användaren, noggrannhet i mätningar, datas ålder osv.

Risken är stor att dolda fel hos data inte uppmärksammas. Egenskaperna hos information framgår i mindre utsträckning av själva data. De data som man betraktar avslöjar inte sin bakgrund om det inte ombesörjs speciellt. Därför är det särskilt viktigt att studera egenskaperna hos information.

En av huvudtankarna med DA och IA är att data ska separeras från tillämpningen och därmed möjligen också från sitt ursprungliga syfte och sammanhang. Det görs i syfte att spara resurser genom återanvändning. Att i denna process fortfarande ha kontroll över datakvaliteten och informationskvaliteten i det som återanvänds är en central fråga.

¹² Ibland kallat "metadata" eller "vetadata".

3 Väv ihop generellt kvalitetsarbete och arbete med informationskvalitet

Kvalitetsarbete generellt och kvalitetsarbete inom området data och information kommer med nödvändighet att vävas ihop alltmer. De arbetsformer och synsätt som vuxit fram inom kvalitetsarbetet såsom ISO-9000 och Total Quality Management kan tillämpas också på data- och informationskvalitet. Kvaliteten på data och information från verksamheten är givetvis också en viktig faktor i ett allmänt kvalitetsarbete. En medveten integration av arbete med informationskvalitet i det generella kvalitetsarbetet kommer att ge en förstärkt verkan totalt sett.

Att informationen och dess kvalitet bör ingå som en del av en generell satsning på kvalitet i en verksamhet är självklart för organisationer vars verksamhet till stor del består av data och informationsarbete.

I enlighet med detta bör en satsning på informationskvalitet inte heller ske isolerat, utan ingå i en bredare satsning på kvalitet inom verksamheten. Det finns flera skäl till det. Dels är verksamhetens kvalitetsmål grunden för mål och krav inom området informationskvalitet. Dels kräver trovärdigheten i en satsning på kvalitet att den satsningen omfattar hela verksamheten och inte bara vissa delar av den. Satsningar på kvalitet, som genomsyrar hela verksamheten, ger grogrunden för den kvalitetskultur som behövs för att satsningar på informationskvalitet ska ge effekt.

I kvalitetsarbetet *i sig* är dessutom data och information viktigt. Det behövs hela tiden uppföljningsdata och mätdata från olika delar av processen och kvaliteten på dessa data måste vara under kontroll.

I arbetet med data- och informationskvalitet är det, precis som i kvalitetsarbete i övrigt, viktigt att ansvaret delas av alla. Ansvaret för kvalitet ligger inte hos kvalitetsorganisationen utan hos var och en. Det är en del i det dagliga arbetet. Kvalitetsfunktionen stöder verksamheten med utbildning, mätningar medan ansvaret för kvaliteten i produkter och tjänster ligger hos frontpersonalen. Var och en som förvaltar och behandlar data har ett ansvar för att bevaka, upprätthålla och säkerställa kvaliteten. Var och en som använder data har ett ansvar för att underrätta sig om kvaliteten eller själv kontrollera den.

Terminologiruta

”**Qualitas** Egenskap, art, beskaffenhet.”

Latinsk och Svensk Ord-Bok Upsala. 1790.

”All through this book, whenever you see the word “quality” read “conformance to requirements”

Crosby 1980

”**Kvalitativ** adj som avser art, inre värde el beskaffenhet **kvalité** *-n -er* s sort beskaffenhet; god beskaffenhet; mest i fråga om varor **kvalitet** *-en -ers*. inre värde; egenskap; sort, beskaffenhet; god beskaffenhet; såsom förled i sammansättning även förstklassig.”

Svenska Akademiens ordlista 1986.

”Kvalitet

alla sammantagna egenskaper hos en produkt som ger dess förmåga att tillfredsställa uttalade eller underförstådda behov.”

Anmärkning (nr 4 av 7):

”Termen »kvalitet» används inte för att uttrycka grad av utmärkthet i jämförande syfte och ej heller i en kvantitativ betydelse vid teknisk utvärdering. I dessa fall ska ett bestämmande adjektiv användas. Till exempel kan man använda termerna

a) »relativ kvalitet» då produkten rangordnas på en relativ skala.

b) »kvalitetsnivå» och »kvalitetsmått» i de fall tekniska bestämningar görs kvantitativt.”

Svensk standard SIS SS 02 01 04 Kvalitet – Terminologi 1987

3.1 Teman inom kvalitetsarbete som kan vara intressanta i samband med informationskvalitet

I diskussioner kring kvalitet brukar ett antal huvudteman förekomma. Vi ställer här frågan om hur de kan vävas samman med frågan om informationskvalitet.

*Kundorientering*¹³. Inriktningen innebär att man har fokus på kundernas behov och anpassar produkter och tjänster för dessa. Detta i kontrast mot en tidigare koncentration på produktionsprocesser med stora serier av produkter som tänks ut på kammaren och sedan hårdlanseras.

Kundorientering kan när det förekommer i sammanhang med data och information tolkas som fokus på användaren. Denna fokusering kan ske på flera sätt. Användaren får en starkare ställning, uppträder mer som en aktör som utnyttjar informationstjänster med känd kvalitet i stället för att vara mottagare av ett dataflöde eller en styrd operatör, en kugge i maskineriet. En annan ansats som anknyter till detta och tas upp längre fram i skriften är användarforum.

Ytterligare en aspekt på detta är att insatser med DA-IA ger möjlighet att utgå från de grunddata man har och sedan göra anpassningar till olika användares varierande behov.

Gör rätt från början. Detta ingår i begreppet Total Quality Management – TQM. Det innebär att kvalitetssystem byggs upp så att fel inte uppkommer och alltså inte behöver rättas. Inom varuproduktion betyder det att man undviker att förbruka stora resurser dels på att producera en viss andel defekta varor och dels på att sedan sortera ut dessa med hjälp av utförliga och dyra kontroller. Kontroller av inkommande material kan sparas in om underleverantören kan garantera noll fel i sina leveranser. Garantikostnader och dåligt rykte på grund av defekta varor som slunkit igenom kontrollen och ändå levererats minskar. Ofta används statistiska metoder för att lokalisera felkällorna så att felorsaker kan åtgärdas.

”Just in time”-filosofin är också en del av detta tänkande. ”Just in time” innebär bland annat att man beställer endast precis så mycket som behövs för att undvika lagerkostnader. Grundantagandet är då att kvaliteten är fullgod till hundra procent, eftersom man inte beställer några extra exemplar för säkerhets skull. Skulle defekter trots allt upptäckas i en sändning så behöver man inte kassera eller kontrollera så stor mängd och det är lättare att finna felkällan.

¹³ För en diskussion kring informationens betydelse för kundorientering i verksamheter se TRIAD-rapport V6, ”Det kundeffektiva företaget”

Detta tänkande härrör dock från kvalitetsarbete i tillverkande industri med stora serier av mer eller mindre identiska och färdigkonstruerade produkter.

Att göra rätt från början i insamling, behandling osv av data kan sägas vara än viktigare, eftersom data ofta inte kan kontrolleras på något annat sätt än genom att man samlar in dem igen från början. En felaktig uppgift kan inte ersättas av en annan uppgift i samma meddelande eller rapport.

Det är dessutom ofta omöjligt att samla in samma data igen. Verkligheten har hunnit ändra sig och mätningen ger ett nytt värde.

I informationssammanhang skulle "just in time" t ex kunna motsvara att man hämtar enskilda uppgifter från en central databas just när de behövs – on-line – i stället för att hämta hem kopior av register och lagra dem lokalt för framtida behov.

En organisatorisk ansats. Det handlar om att bygga upp procedurer, fördela ansvar, införa ett beständigt system för "quality management", upprätta kvalitetspolicy osv för att fastställa och uppnå avsedda kvalitetsnivåer. ISO-9000 har till stor del en sådan organisatorisk ansats.

Denna ansats är naturligtvis relevant även för information. Organisatoriska åtgärder krävs både för god datakvalitet och informationskvalitet. Generellt kommer organisatoriska satsningar för kvalitet även att stödja en god informationskvalitet.

IA/DA-organisationen har givetvis en viktig roll här och är en av de första som bör arbeta enligt de kvalitetsprocedurer som tas fram inom organisationen.

Betonande av personalens kompetens och motivation. Detta tas upp i ISO-9000 serien och på andra håll. I fråga om informationskvalitet kan man se betonandet av personalens kompetens och motivation som ytterligare en påminnelse om användarens eller aktörens centrala position. Det är användaren och aktören som ombildar data till information och det är effekten av hans eller hennes handlande som utgör den slutliga prövningen av kvaliteten. Kunskaper om ämnesområdet, informationssystemet, osv är lika avgörande för resultatet som t ex korrektheten i data.

Informationens betydelse. Att data och information är viktigt för kvalitetsarbetet är kanske alltför självklart för att det ska uppmärksammas som en egen huvudfaktor. Det nämns dock på flera håll. En vanlig uppmaning i samband med kvalitetsarbete är att man ska mäta – och börja mäta tidigt. Kvalitetsmått och -indikatorer ska identifieras och mätningarna ska påbörjas innan det operativa kvalitetsarbetet startas. Då kan man se effekterna av de insatser som görs. I anvisningarna för *Utmärkelsen Svensk Kvalitet* (Institutet för Kvalitetsutveckling 93) betonas på flera ställen informationens betydelse för kvalitetsarbetet i form av data och fakta som samlas in. Att dessa data måste hålla en specificerad och känd kvalitet är också givet.

Kulturförändring, särskilt med ett starkt engagemang hos ledningen. Det betonas ofta att det krävs en stor omsvängning i tänkande och beteende i organisationen. Särskilt gäller detta ledningen. Ledningen måste t ex enligt Peters (Peters 88) vara "nästan fanatisk" i sin strävan att höja kvaliteten.

Kvalitet är lönsamt. Man hävdar att kvalitetsarbete inte är en kostnad för företag utan en ren vinst. Philip B. Crosby, tidigare kvalitetschef på ITT, har titeln "Quality is free" på en bok (Crosby 1980). Han menar att det inte är kvalitetsarbete som kostar utan kvalitetsbristerna – alla kostnader för att inte göra rätt från början. Kvalitetsarbete ger lönsamhet genom mindre kostnader för fel, spill och kontroller. Det ger högre intäkter genom man behåller kunder, får fler kunder och mer nöjda kunder. Man hävdar också att optimeringstänkande enligt det traditionella diagrammet där kostnadskurvorna för kvalitetsbrister och förebyggande åtgärder möts i en viss optimeringspunkt är falskt: "Det finns alltid ett sätt att få en bättre kvalitet till en lägre kostnad" (Bergman och Klevsjö 1991). En aspekt på detta är att det kan vara svårt att fånga in alla samband i en kalkyl och bedöma sannolikheter. Ett enda fel kan förstöra ett företags goda rykte. Räknar man med att det kan hända, så motsvarar kostnaden för ett avhjälp eller förebyggt fel alltid värdet av företagets överlevnad. Som tidigare nämnts måste dock krav på kvalitetsnivåer ha specificerats för att fel och brister ska kunna identifieras. I en del fall är rätt eller fel självklart, i andra en fråga om gränsvärden och syfte eller ändamål

Fortlöpande förbättring – Kaizen. Kaizen är ett japanskt uttryck som står för ständig förbättring. Man betonar att kvalitetsarbetet är en ständigt pågående process med små steg, inte en stor satsning av engångskaraktär. Det innebär också att man fortsätter att göra förbättringar, även när inga problem och fel finns att rätta till och allt fungerar bra. För informationssystem betyder detta att de byggs så att en ständig utveckling i små steg är möjlig – systemevolution snarare än revolution där helt nya system introduceras med långa mellanrum, medan utvecklingen däremellan står stilla. Systemkrav i en tidigare generation får inte låsa in organisationen och omöjliggöra förändringar inom systemet. Kaizen, de ständiga och stegvisa förbättringarna gäller givetvis också informationskvaliteten.

3.2 Praktisk vägledning inom kvalitetsområdet

Det finns en rikhaltig litteratur inom området kvalitet. Från denna kan uppslag hämtas, även om specifika beskrivningar av just informationskvalitet som regel inte ingår.

ISO-9000-serien ger vägledning i kvalitetsarbetets procedurer och former. Här kan man t ex få anvisningar för hur ett kvalitetssystem byggs upp. Detta kan givetvis tillämpas för att få en lämplig form på arbetet med informations kvalitet. Den del av ISO-9000-serien som är mest tillämplig är troligen SS-ISO 9004: "Kvalitetssystem Del 2: Allmänna riktlinjer för tjänster". Övriga handledningar och litteratur för kvalitetsarbete som går utöver ambitionsnivån i ISO-9000 kan också ge vägledning för arbetets former. Det handlar dock även där om vägledning – inte konkreta beskrivningar om hur man tacklar de specifika problemen med informationskvalitet.

SS-ISO-9000-3 behandlar "riktlinjer för tillämpning av SS-ISO 9001 vid utveckling, leverans och underhåll av programvara". Denna standard har även inkluderats i sin helhet i TickIT-handboken (SFK-Data, VI, 1993) och ger ytterligare handledning för kvalitetsarbete vid konstruktion av programvara och kvalitetsrevision. Programvaran är en viktig länk i den kedja av faktorer som påverkar informationskvaliteten. Tillämpning av dessa riktlinjer kan därför mycket väl vara ett stöd. Det behövs dock aktiviteter med en mer specifik inriktning på data och information i sig och i Tick-IT ges inte några utförliga anvisningar för det.

En vägledning i ämnets terminologi ges i Rapport ITS 6 "Terminologi för informations säkerhet", mars 1994. En terminologi med termer och definitioner är, som alla DA-intresserade redan vet, ett viktigt verktyg för framgångsrikt arbete inom ett område. Terminologin är till stöd t ex när kvalitetskrav ska fastställas eller faktisk kvalitet redovisas.

RRV-rapporterna "Rätt data?" och "Fel data kostar!" har också bilagor som ger metodvägledning. Rapporterna finns i skrivande stund fortfarande i lager och kan beställas från Riksrevisionsverket.

4 Resultat med enkla åtgärder

Det är möjligt att med relativt enkla åtgärder börja ta sig an området informationskvalitet. Ett exempel är *användarforum*.

Användarforum används här som en samlande beteckning på olika slags möten med informationsanvändare som har till syfte att:

- ta reda på hur data används
- finna fel och brister i det nuvarande läget
- undersöka vad som behövs för att kunna fatta bra beslut
- ta fram förbättringsförslag

Direktkontakten mellan den som arbetar med datasystem och användare eller aktörer är central. Det är möjligt att under enkla former åstadkomma resultat.

Vid Trafikregistret inom Vägverket gjordes en endagsseminarie som innebar att all personal medverkade i grupparbeten kring informationskvalitet. Man började med ett cirka en timme långt föredrag vars innehåll i huvudsak var baserat på RRV:s "Rätt data?" (RRV 90). Det hela var upplagt på ett underhållande sätt, med tidningsartiklar m m som exempel.

Föredraget var engagerande för personalen och gav en tydlig bild av vad informationskvalitet handlar om. Före denna genomgång hade man olika tolkningar om vad informationskvalitet är eller kan vara. Efter föredraget arbetade mindre grupper med frågor. Uppgiften var att inventera svaga punkter, felkällor, brister i kompetens och att ta fram förändringsförslag. Resultatet blev ett antal sidor med iakttagelser och förslag.

En sådan aktivitet kan förväntas leda till uppslag, inte bara i fråga om informationskvalitet utan också när det gäller andra förändrings- och förbättringsförslag.

En utförlig och praktisk beskrivning på ett arbetssätt med inriktningen användarforum ges i "Att mäta informationens användbarhet" (Ledell 93). Skriften beskriver hur en grupp på högst 15 personer med en handledares hjälp och med lappteknik analyserar en informationsmängd. Analysen delas in i områdena riktighet, instruktioner, tillgänglighet, kunskap, layout och övrigt. Ett antal olika aspekter täcks alltså. Förslag till åtgärder tas också fram av gruppen. I skriften finns även underlag för presentation av ämnet med förlagor till OH-bilder m m.

I detta sammanhang bör dock en fråga om perspektiv uppmärksammas. En aktivitet av detta slag kan tolkas som att man från ADB-sidan på nåder bjuder in användare så att de kan framföra sina synpunkter och att man sedan drar sig tillbaka för att besluta om vilka förändringsförslag som eventuellt ska bifallas. När ett sådant inflytandeförhållande råder är det mycket centralt att man har en klart uttalad serviceinställning.

Utgångsläget bör vara att ansvaret för informationen finns *i verksamheten*. Användarna/aktörerna och de kvalitetsansvariga tar då initiativ till aktiviteten och ställer i sin beställarroll krav på systemsidan. Detta kräver dock en medvetenhet och kunskap om data- och informationsfrågor som kan behöva byggas upp med utbildning eller på annat sätt.

Användarna/aktörerna känner dock inte till, och kan alltså inte upptäcka alla fel. Kontakt med användarna är alltså en nödvändig – men inte tillräcklig – aktivitet för att man ska kunna kontrollera informationskvaliteten.

Användarkontakterna kan dock ge värdefulla uppslag till områden där man bör söka vidare efter fel och brister, eller peka på kvalitetsnivåer som inte alls är motiverade.

5 Ställa krav på informationskvaliteten

”Vi måste ha försäljningssiffrorna *snabbt*”, sa Åke, försäljningschefen. ”Det gäller att kunna agera fort när någonting händer på marknaden. Och då krävs *aktuell* information”.

”Jaa i och för sig... men den måste vara korrekt”, sa ekonomichefen, ”*Inte en siffra* lämnar min avdelning utan att den är ordentligt kontrollerad. Och det tar tid, det får du acceptera!” sade han och stirrade stint på Åke.

Åke, som länge irriterat sig på ”den där kamrerstypen som ligger så bra till hos VD”, blev röd i ansiktet men svarade med tillkämpad vänlighet.

”Ja, jag får väl göra det”.

På vägen ut mumlade han

”Jag bygger ett eget system. Det är inget annat att göra, Svenne kan nog fixa nånting.”

Är detta en situation som känns igen? Man kanske kan undra var kopplingen till informationskvalitet finns. Avvägning mellan olika kvalitetsfaktorer vid målkonflikter är dock ett exempel på en av de frågor som måste tacklas i kravanalysen.

När det står klart att en medveten kravställning behövs, stiger troligen vissa områden fram som mer angelägna att ta itu med. Här ges en översikt över aspekter som kan vara aktuella att beakta när krav ska fastställas på ett medvetet och ordnat sätt: Vilka aspekter ska tas med i de enskilda fallen? Med vilken utförlighet ska de behandlas? Hur ska de anpassas från fall till fall med hänsyn till de specifika förutsättningarna?

I avsnittet ges också en fortsatt och något mer utförlig behandling av frågan om vad informationskvalitet är samt en kortare genomgång av aktuell terminologi.

Kravspecificeringen som process

Kravspecificeringen kan riktas mot kunden/användaren eller mot systemet/processen. Inriktningarna behöver inte stå i motsats till varandra utan bör snarare komplettera varandra.

Kund – användarinriktningen går ut på att anpassa information till mottagarens krav och förutsätter att mottagaren har det fulla ansvaret och befogenheter att fastställa kraven. Mottagaren med sina behov och sin användningssituation osv blir därför central.

Arbetet med att på detta sätt fastställa kvalitetskraven på information har stora likheter med att specificera kvalitet hos tjänster. Tjänsten utformas stegvis i ett samspel mellan utförare och mottagare och detta samspel fortsätter fram till den tidpunkt då tjänsten har utförts och man kan utvärdera om mottagarens önskningar och behov är uppfyllda. Kvaliteten i denna process blir alltså en del av informationskvalitetens Total Quality Management – TQM.

Även användaren/mottagaren har en stor betydelse för slutkvaliteten. Det gäller t ex den del där användaren utövar sin aktörsroll genom att påverka ett uttag ur en databas eller presentationen av data. Naturligtvis har det också betydelse vilken kunskap användaren har i övrigt.

För att kunna ställa krav på data- och informationskvalitet måste kravställarna uppfylla vissa förutsättningar.

Exempel på vad som krävs är:

- kunskap om hur informationssystem fungerar och vilka risker som finns
- kunskap om sambanden mellan informationskvalitet och verksamhetens kvalitet och produktivitet
- terminologi för att formulera krav
- olika former av teknik för att mäta data- och informationskvalitet

Om ingen är medveten om vilka problem som orsakas av bristande kvalitet och om ingen har ambitionen att åstadkomma kvalitet i verksamheten, då saknas även incitamentet och därmed en viktig förutsättning för kravställande.

Med en inriktning på *systemet/processen* ser man längre än till användarna eller mottagarna. Man tar då också hänsyn dels till att mottagaren inte alltid har det fulla ansvaret och alla befogenheter som krävs, dels till effekterna av de kvalitetsbrister som kan uppstå utom synhåll för informationsanvändaren. Här tillämpar man ett arbetssätt där data följs från ax till limpa; från det att de föds till dess att effekterna av användningen kan ses. Till detta arbete hör att identifiera system- och ansvarsgränser och kartlägga vilka informationsprocesser eller -kedjor som är berörda.

Krav på vad? – Mer om vad data- och informationskvalitet är

I detta avsnitt går vi vidare i frågan om vad data- och informationskvalitet är och relaterar till en del användbar terminologi.

Uppdelningen i "data i sig och data för mig" från tidigare avsnitt kan relateras till skillnaden mellan datakvalitet och informationskvalitet.

Följande distinktion brukar göras mellan data och information:

Terminologiruta

"Data

Representation av fakta, begrepp, eller *instruktioner* i form lämpad för *överföring*, tolkning eller bearbetning utförd av människor eller automatiska hjälpmedel."

"Information

Den innebörd en mottagare lägger in i givna data"

(SEN 01 16 01)

Det bör då även vara möjligt att skilja mellan datakvalitet och informationskvalitet.

Terminologiruta

”Kvalitet handlar om egenskaper. Informationskvalitet är sålunda egenskaper hos information.”

”Man bör skilja på data respektive informationskvalitet. Datakvalitet är relaterad till *representationen* av fakta etc. (...) Informationens kvalitet är relaterad till *tolkningen* eller innebörden av data. När vi talar om informationskvalitet måste vi därför ta hänsyn till hur erhållen information, dvs tolkade data, påverkar användaren, det berörda objektet samt omgivningen.”

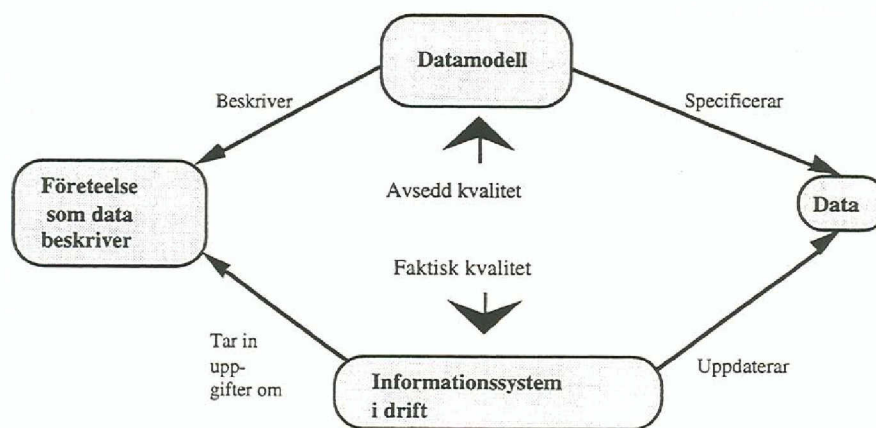
”**Datakvalitet** egenskap hos data som behandlats enligt specifikation och inte förändrats på något annat sätt. Datakvalitet är en sammanfattande benämning på egenskaper hos data i sig.”

”**Informationskvalitet** användbarhet, också med hänsyn till dess effekt, av information för en given användare och ett givet problem.”

(ITS 94)

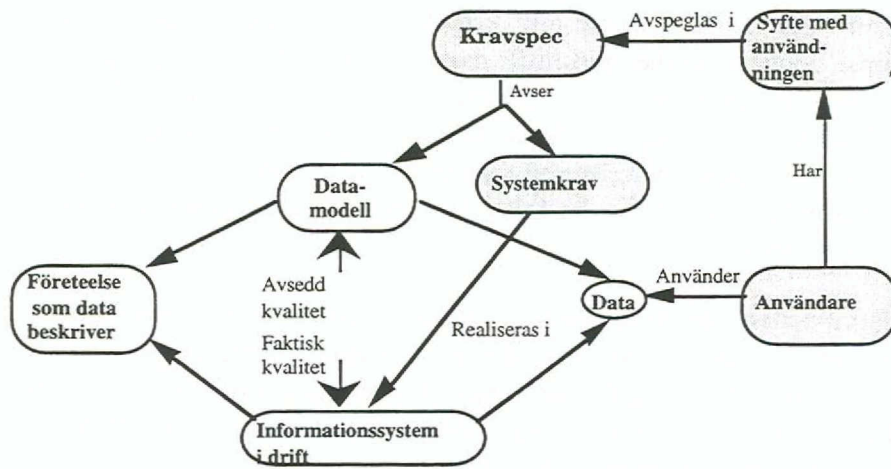
Två vägar till data- och informationskvalitet

Den kvalitet som data och information får, uppkommer huvudsakligen på två sätt: dels genom den avbildningsmodell (datamodell) man har valt och dels genom det system som uppdaterar data utifrån förändringar i objektet (se figuren nedan). Modellen och tillhörande definitioner, beskrivningar m m specificerar den *avsedda kvaliteten*.



Figur 2

Datakvaliteten är beroende av kvaliteten på datamodellen och övriga systemspecifikationer. Därmed är datakvaliteten också beroende av det arbete som ligger bakom produkterna. Datakvaliteten är också beroende av det existerande systemet som hanterar de uppdateringar som görs. Systemets funktion kan avvika från den kvalitet man avsåg att uppnå – både permanent och tillfälligt. Resultatet blir den *faktiska kvaliteten* hos de data som används.



Figur 3

Ur datamodellen och övriga specifikationer kan man utläsa den avsedda kvaliteten. Det är alltså grunden för vad som kan uppnås *maximalt*, både av specifik kvalitetsnivå och klarhet om vad den avsedda kvaliteten är.

Kvaliteten i arbetet med datamodellen och definitionerna ligger bland annat i hur väl man lyckas avbilda de väsentliga egenskaperna hos objektet, dvs de som är väsentliga i förhållande till syftet.¹⁴ Detta arbete kan ses som ett konstruktionsarbete där data och information ska konstrueras med utgångspunkt i verksamhetens krav. Datamodellerna har sin grund i verksamhetsmodeller. Precis som system måste verksamhetsmodeller, datamodeller och informationsmodeller bli föremål för en sammanhållen "konstruktion". Enligt rådande praxis ingår också vid denna konstruktion att ha en kontrollerad versionshantering och att upprätthålla konsistensen mellan de konstruktioner som görs för olika tillämpningar.

God kvalitet i DA-arbetet är alltså en av de hörnpelare som håller upp data- och informationskvaliteten.

¹⁴ Ämnet modellkvalitet behandlas i en egen TRIAD-rapport (SISU 93).

Utifrån syftet väljer man *vad* som ska avbildas i modeller och data – och även *hur* det ska göras. Det som kan avbildas i modellen kan vara företeelser, entiteter, objekt, egenskaper, attribut, siffror, text, fast nomenklatur, bild, ljud osv.

Kopplingen till syftet innebär en subjektivitet i urval och representationsform och denna subjektivitet är viktig att uppmärksamma. Om man t ex vill samla data om en person finns det ett otal egenskaper att välja mellan. Syftet avgör om det är inkomst, utbildning, blodvärden, hattnummer, politisk uppfattning eller midjemått man inriktar sin datainsamling på, och hur man beskriver dessa egenskaper. I idealfallet tar man vid datamodelleringen också upp ytterligare kvalitetsaspekter och fastställer nivåer för t ex feltolerans (inklusive vad som är rätt respektive fel), noggrannhet vid mätningar osv, så att det finns fasta krav att utgå från vid systemkonstruktionen.

Informationssystemets egenskaper vid drift avgör sedan hur den *faktiska* kvaliteten blir, dvs hur pass nära den *avsedda* eller specificerade kvaliteten man kommer när det gäller de uppgifter som lagras i registren. Fel och brister samt svagheter i systemet leder till att data kan komma att klassificeras fel, registreras och kodas fel, ändras oavsiktligt eller på ett annat sätt än avsett. Data kanske aldrig kommer in, fördröjs i sin förflyttning genom systemet, inte är tillgängliga för användaren osv. Uppgiftslämnare kan lämna falska uppgifter eller underlåta att lämna uppgifter. Ett känt exempel är att privatpersoner mantalsskriver sig på en adress och bor på en annan. Detta har med datas validitet att göra.

Man kan då tala om registerkvalitet som kan mätas eller bedömas i en blixtbild (tillstånd), eller i en bild av data i sig, av registret i sig (motsvarar återigen på ett ungefär datakvaliteten).

Denna kvalitet är *densamma* oavsett lagringssätt och system: tape, skiva, kortlåda, CD, papperslista.

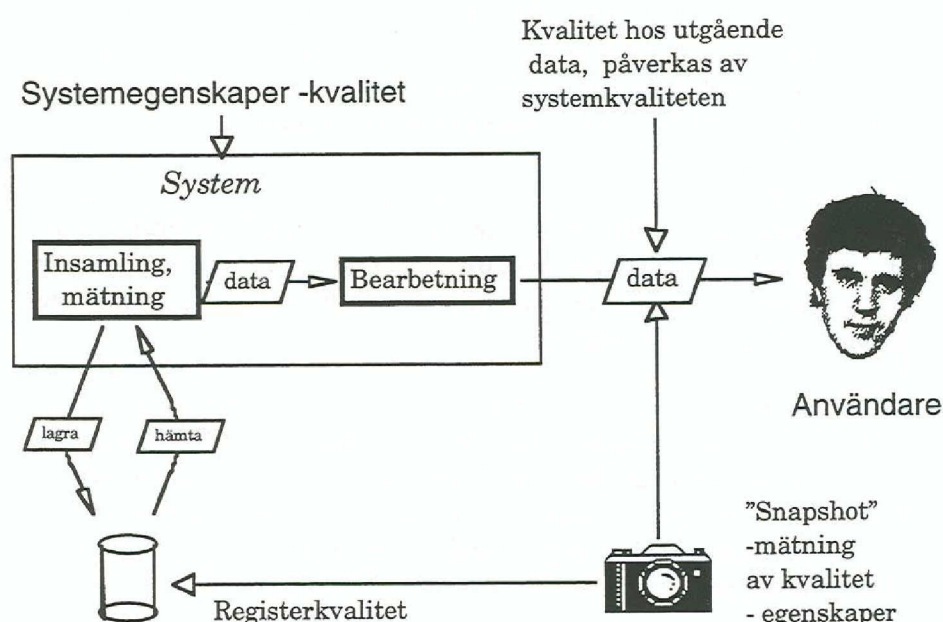
Bilden kan tas var som helst i flödet inom ett system.



Kameran kan illustrera ett "snapshot" av kvaliteten på data på en viss plats vid en viss tid.

Ett exempel på kvaliteten kan vara antalet rena felaktigheter (objektiva fel) i de data som man tar bilder på, t ex telefonnummer med omkastade siffror i.

Kvaliteten hos informationstjänsten (funktionen) ur användarens synvinkel kan vara t ex ledtider, tillgänglighet, presentation eller bearbetningsmöjligheter. Samma grunddata kan presenteras på olika sätt av en informations tjänst eller funktion. Detta påverkar den slutliga informationskvaliteten t ex i form av användarens tolkning av data. Denna tolkning kallas också informationsbildning. Tjänstens tillgänglighet i olika avseenden och olika slags stöd och service som kan erhållas har också betydelse.



Figur 4

I systemkvaliteten kan man möjligen också inkludera aspekten att informations kvalitet handlar om att få precis den information som just jag vill ha och behöver, i en form som passar mig. Det handlar därmed också om att *inte* behöva ta del av stora mängder data som inte är intressanta alls – "inte behöva kamma igenom stora datahöstackar sökande efter nålen". Det handlar också om att jag kan få informationen presenterad i en form som passar just mig.

Faktisk kvalitet kan ibland föredras framför avsedd kvalitet

Avvikelser från den kvalitetsnivå som var avsedd från början behöver inte alltid innebära en försämring i förhållande till syftet. När systemet väl är i drift, kan det visa sig att man tänkt fel på vissa punkter, att kraven ändrats eller från början missuppfattats. Detta aktualiserar en ny kvalitetsaspekt i systemet: i vilken utsträckning det kan ändras för att ge en annan kvalitet än vad som först var avsett. Detta blir då en ny avsedd kvalitet.

Arbete enligt DA-principer kan ibland ge en avsedd kvalitet som är mindre entydigt definierad. I stället för en entydig kvalitetspunkt har man snarare en större yta eller möjlighetsrymd där nya syften kan tillgodoses. Det är då viktigt att gå igenom möjliga syften i förväg.

Ändringar i system måste vara kända

Viktigt är dock att den som ansvarar för systemet vet vilka ändringar som görs och även informerar alla intressenter om detta. Detta är en del av arbetet med *deklarationer*. Om det inte görs kan kvaliteten hamna helt utanför kontroll och vara okänd.

Datafält kan t ex helt eller delvis användas på ett annat sätt än som var tänkt från början. I själva verksamheten kan man hantera detta. Om data från systemet också ska rapporteras vidare, t ex för statistiska ändamål eller för resultatuppföljning, kan effekterna däremot bli konstiga.

Ett exempel är ett system för hantering av ekonomiskt bistånd till kommuninnevånare. När ett ärende togs ur systemet för manuell handläggning lades en biståndssumma på 1 000 000 dkr in som en signal för undantagshantering. I statistiken kom denna kommun högt i fråga om ekonomiskt bistånd (Brammer 93).

Ett annat exempel är när ett kodnummer i ett system återanvänds i en ny betydelse eller när nya koder införs i ett existerande system. Om gamla data läses med den nya koden kommer de att tolkas på ett nytt sätt. Journal bör föras om sådana förändringar för att säkra spårbarhet (Brammer 93). Denna journalföring kan ingå i datakataloger eller i förteckningar över dataelementtyper.

Andra aspekter på informationskvalitet

Det som beskrivits ovan är aspekter på data och informationskvalitet som i första hand är kopplade till traditionell databehandling, dvs behandling av strukturerade data. Det finns också andra aspekter på ämnet. Rapporten går inte närmare in på dem men ett par exempel ges nedan.

Kvalitet i information handlar om hur en rik bild kan åstadkommas

En del av kritiken mot ADB-system som stöd i arbetet handlar om att arbetsinnehåll och informationsrikedom reduceras mycket starkt vid datorisering. Den informationsrikedom som direktkontakt ger försvinner. Man brukar bland annat säga att det som kallas för tyst kunskap reduceras. Med tyst kunskap menas det som kan läras in genom egen erfarenhet eller genom att man iakttar mycket kunniga personer i arbete, men som inte lätt kan formuleras på ett entydigt sätt och därmed läggas in i ett traditionellt ADB-system.

De nya möjligheter som hantering av bild och ljud, multimediateknik, videokonferenser, virtuell verklighet osv ger, skulle möjligen till någon del kunna råda bot på de nackdelar med reduktion av informationsrikedom som kodning av data innebär. Det skulle i vissa fall inte vara nödvändigt att gå omvägen via klassificering, kodning osv av data. Grunddata, t ex en bild, skulle kunna föras över direkt. Exempel finns på hur läkare via kommunikationsnät med hög kapacitet snabbt kan delge en kollega på annan ort en röntgenbild och få ett utlåtande som styrker eller kompletterar den egna diagnosen (Teldok 94).

Med ständigt lägre kostnader för lagrings- och behandlingskapacitet, skulle det finnas möjligheter att registrera och lagra en mycket stor mängd data för varje objekt. Presentationen skulle sedan kunna ske på olika sätt. Samma sak kan presenteras på olika sätt och jämförelser kan göras.

Här är vi dock inne på ett svårgripbart område. Med en rik bild vet vi inte vilka data som användaren tolkar och hur. Endast en liten del av de sinnesintryck som tas emot bearbetas på en medveten nivå i hjärnan¹⁵, och användaren kommer inte alltid att kunna redogöra för vilka data som exempelvis ett beslut baseras på. Värderingen av informationskvaliteten kan då bara göras genom att man ser på effekterna av mottagarens handlande. Då finns å andra sidan problemet att andra faktorer som vi inte känner till kan ha spelat in. I situationer där viktiga beslut ska fattas kommer man säkerligen även i fortsättningen att behöva redovisa entydiga beslutsunderlag. Att hitta lämplig användning för de nya tekniska möjligheterna är troligen under rätt så lång tid en uppgift för forsknings- och utvecklingsprojekt. Det ger dock nya intressanta perspektiv på ämnet informationskvalitet!

¹⁵ Tor Nørretranders skriver i sin bok "Märk världen" att medvetandets kapacitet är cirka 40 bit/s och att endast en miljondel av det som våra sinnen uppfattar når vårt medvetande. Hjärnan i övrigt tar dock in alla sinnesdata som strömmar in och lagrar dem troligtvis till större eller mindre del. Man har t ex kunnat visa att människor i drömmen kan återkalla subliminala stimuli (som ligger under den gräns som kan uppmärksammas av medvetandet, t ex en inklippt enstaka reklamruta på 1/25 sek i en film) som de exponerats för i vaket tillstånd.

Kvantitativ eller kvalitativ metod.

En annan aspekt på detta är skillnaden mellan kvantitativa och kvalitativa data, eller kvantitativ och kvalitativ metod. Den kvantitativa metoden är enligt Eneroth (Eneroth 92) en undersökningsmetod som "...syftar till att fastställa en kvantitet, dvs mäta hur mycket som finns av något (...) medan en kvalitativ metod söker efter den sorts kvalitet som är typisk för en företeelse eller som skiljer den från andra, menar den kvantitativa metoden alltså att det typiska eller skiljaktiga finns i mängden av en viss kvalitet". Vi kan jämföra olika personer, t ex kvantitativt, genom att jämföra inkomst, ålder och antal barn, eller kvalitativt genom att i det ena fallet beskriva t ex framgångar i yrkeslivet och i det andra fallet berätta om strapatsrika resor på fritiden. I både kvantitativ och kvalitativ metod använder man data, men de väljs och behandlas på olika sätt. I fråga om ADB-stöd, kan man kanske säga att relationsdatabaser¹⁶ passar bäst ihop med den kvantitativa ansatsen, medan teknik för multimedia och s k hypertext samt eventuellt objektorientering bättre beskriver kvalitativa egenskaper.

Terminologin är ett verktyg för kravställningen

För att kunna precisera kraven och även för att kunna redovisa avsedd och faktisk kvalitet, behövs en terminologi som går in mer på detaljer än huvudindelningen i validitet och relevans. Under årens lopp har vissa huvudbegrepp för data- och informationskvalitet vuxit fram. De termer som används i praktiken varierar dock liksom den betydelse de ges. Aktualitet, noggrannhet och detaljeringsgrad är troligen de vanligaste termerna, eventuellt också korrekthet. Korrekthet används dock i vissa fall som synonymt med validitet, dvs som en beteckning för ett bredare begrepp.

Användningen av terminologin kan variera beroende på det syfte som användaren har. En användare/aktör är kanske inte intresserad av vad felen i ett register beror på, utan vill bara veta hur många och stora de är. I en deklARATION som den användaren ska läsa är det då kanske inte meningsfullt att redovisa fel som beror på bristande aktualitet för sig, registreringsfel för sig, mätfel för sig osv. En person vars syfte är att finna åtgärder för att förebygga felen, har dock givetvis stor nytta av en sådan särredovisning.

¹⁶ I en relationstabell beskrivs som regel ett objekt med ett antal egenskaper, samma egenskaper för alla objekt. I och för sig måste inte alla egenskaper beskrivas *kvantitativt* och det bör även vara möjligt att låta bli att registrera vissa egenskaper för vissa objekt. Det senare skulle dock kunna betraktas som en brist i registrets fullständighet.

Otillräcklig standardisering i fråga om terminologi, mätmetoder och skalor för data- och informationskvalitet nämns som ett problem av Brammer (Brammer 93), och han föreslår ett standardiseringsarbete inom CEN/TC311.

I Sverige finns en hel del termer definierade i Rapport ITS 6¹⁷. ITS 6 är inte en fastslagen standard, men väl ett hjälpmedel för att lösa terminologi-frågor. I praktiken används, åtminstone på en del håll, delvis andra termer än de som finns i ITS 6. Om terminologin ska inkorporeras i databas- och datakatalogverktyg eller IRDS, behövs kanske striktare standardisering på internationell nivå och mer detaljerade anvisningar för hur kvalitet ska deklarerars. Men innan en sådan standard fastställs, vore det troligen värdefullt att få erfarenheter av att deklarerars kvalitet enligt de rekommendationer som nu kan ges redovisade. Enligt RRV:s studie är kvalitetsdeklARATIONER ovanliga (RRV 1990).

När kvalitetsmått för en viss verksamhet ska redovisas, är det troligen ofta lämpligt att använda en terminologi som knyter nära an till verksamheten. De termer som används blir då kanske mer konkreta aspekter på dem som t ex finns i ITS 6. Om man upptäcker något slags fel som inte alls täcks i den befintliga terminologin, eller ligger i närheten av någon av dess termer, får man skapa en ny. Givetvis med kraven att det nya begreppet ska ha en entydig definition och ett upplysande namn, som inte redan används i någon annan betydelse. Detta spörsmål behandlas också i avsnittet om kvalitetsdeklARATIONER. Terminologiarbete behandlas i TRIAD-projektets modelleringshandbok, delen *Namnsättning* (TRIAD 93).

ITS 6 beskriver, som tidigare nämnts, ett antal termer som är användbara för området. Några exempel därifrån används i den fortsatta texten. Allt är dock inte med här, så den intresserade läsaren uppmanas att skaffa och läsa den nya ITS 6 från 1994.

Tid och aktualitet

Tiden går och världen förändras. Data är en färskvara som åldras, men åldrandet syns inte på data i sig, eftersom åldrandet beror på förändringar i det som data avbildar. I ett längre tidsperspektiv tillkommer förändringar i språkbruk, verksamhet osv som gör tolkningen av data alltmer vansklig. Tiden är ett av de centrala problemen med data- och informationskvalitet.

Aktualiteten hos data är ett relativt begrepp, där en tidpunkt för händelse eller insamling jämförs med nutiden. Den bästa aktualiteten, dvs den lägsta åldern på data som kan uppnås, är beroende av olika ledtider i systemet.

¹⁷ Terminologi för informationssäkerhet.

När kraven på aktualitet och registrering av olika tider ska formuleras, behövs uppgifter om hur snabbt det som data avbildar ändrar sig, om data fortfarande har ett användningsvärde även om de är historiska osv. Vissa data kan bli helt oanvändbara så fort verkligheten har hunnit ändra sig. Andra kan ha full relevans under förutsättning att det är känt vilken tidpunkt de gäller för.

Terminologiruta

”aktualitet ålder hos data

Är av intresse där ålder hos data har betydelse för dess användning. Alla lagrade data är per definition gamla. Det tar alltid en viss tid, **ledtid** från insamling till dess att data är tillgängliga hos användaren.”

”användningstid tidpunkt vid vilken information används

Skillnaden mellan händelsetid och användningstid är åldern på data, **aktualitet**.”

”händelsetid tidpunkt då den händelse man vill mäta inträffade

Data avser denna tidpunkt.”

”ledtid tiden mellan händelsetid och då data finns tillgängliga

Jämför **aktualitet**.”

”mätperiod tidsperiod under vilken mätning ägt rum”

”mättidpunkt tidpunkt på vilken en mätning baserar sig

Mättid kan vara samma som händelsetid.”

”periodicitet hur ofta mätning utförs

Behov av periodicitet beror på hur ofta förändringar sker i objekts egenskap samt med vilken noggrannhet man mäter.”

”tidstäckning total tidsperiod alla mätningar (av en viss variabel) omfattar”

ITS 94

Noggrannhet och detaljeringsgrad

Noggrannhet och detaljeringsgrad har att göra med hur värdena hos objektets egenskaper mäts. Det är viktigt att hålla isär begreppen. *Detaljeringsgraden* har att göra med mätskalan, hur små eller stora skillnader som kan mätas. Klassifikationsscheman – nomenklaturer (som det tidigare nämnda SNI) är också skalor. De kan vara nivåindelade så att det är möjligt att välja detaljeringsgrad när man använder en nomenklatur. *Noggrannhet* handlar om hur nära ett idealvärde man har kommit i en mätning, dvs om det finns ett mätfel och hur stort detta i så fall är. Samband mellan detaljeringsgrad och noggrannhet finns också. Om man vill åstadkomma en relativt sett högre noggrannhet så krävs en mer detaljerad skala för mätningen. Om man mäter med en mycket grov skala så kan större avvikelser hos objektet förekomma utan att något mätfel uppstår, dvs noggrannheten är paradoxalt nog god.

Termen detaljeringsgrad bör även kunna användas för att beteckna teknisk kvalitet, t ex i form av upplösning i punktgrafik. Den ursprungliga upplösningen hos en digitaliserad bild avgör den slutliga kvaliteten, t ex efter att den digitaliserats om några gånger vid faxöverföringar. En annan situation där den tekniska kvaliteten är viktig, är när en text som skickats med fax ska återomvandlas till teckenkoder. Detta kan i bland vara enklare än att försöka föra över en textfil på annat sätt. Då kommer den tekniska kvaliteten, upplösningen, att ha betydelse för resultatet.

Terminologiruta

”**noggrannhet** hur väl de specificerade egenskaperna mäts och erhåller ett värde.

Mätningen sker med en viss noggrannhet enligt en skala (enligt nominal- ordinal- intervall- eller kvot- skalan) som kan vara mer eller mindre detaljerad. Mätfel kan t ex anges genom mått på standardavvikelse.”

”**detaljeringsgrad** grad av upplösning.

Är beroende av bl a mätskalan. Kan t ex uttryckas i hur många klasser som beskriver ett fenomen. Ju fler klasser desto högre detaljeringsgrad. Detaljeringsgrad i aktualitet är periodicitet. En skala med hög detaljeringsgrad, t ex många klasser, innebär en större risk för fel. Jämför **periodicitet**.”

ITS 94

Täckning och bortfall

Täckning och bortfall har att göra med om ett register innehåller alla uppgifter som det ska. Dessutom har det att göra med om alla uppgifter är med för alla objekt som finns registrerade. Om objekt som inte ska vara med finns med, föreligger övertäckning. Om å andra sidan objekt saknas, föreligger undertäckning. En registerdefinition ska ge ett klart svar på frågan om vad som ska ingå i registret.

Terminologiruta

”**registerdefinition** definition i vilken avgörs vilka (typer av) objekt som ska ingå i ett register”

”**täckning** utsträckning i vilken ett register innehåller uppgifter om exakt de objekt som det enligt **registerdefinitionen** skall innehålla”

”**bortfall** tillstånd när någon egenskap eller variabel saknas hos ett i systemet avbildat objekt”

ITS 94

Objektiva fel

Fel som inte kan beskrivas korrekt med de tidigare redovisade termerna behöver en egen term. Det gäller åtminstone när felorsaker och förebyggande av fel avhandlas.

Terminologiruta

”**objektivt fel** fel som inte är aktualitets-, täcknings- eller noggrannhetsfel.”

”**objektiv felfrihet**: tillstånd som innebär att inga fel inträffar i processen som påverkar informationskvaliteten.

Fel kan inträffa i hela flödet från informationsinsamling till användning. Fel kan vara systematiska eller osystematiska. Systematiska fel är fel orsakade av en eller flera faktorer som påverkar enligt någon lag eller mönster. Vid upprepad mätning kommer samma systematiska fel att uppstå. Osystematiska fel är orsakade av slump, slarv eller dylikt.”

”**feltolerans** Mått på hur mycket fel som kan tolereras innan erhållen information är oanvändbar.”

Med fel avses här andra fel än de som uppstår på grund av brister i aktualitet, täckning och noggrannhet, dvs objektiva fel. Data som faller utanför den accepterade feltoleransen kan ge negativa effekter vid viss användning.”

ITS 94

Termen **jämförbarhet** finns inte definierad eller särskilt kommenterad i ITS 6. Jämförbarhet kan inte beskrivas för *en* datamängd. Datamängden måste jämföras med en annan för att man ska kunna säga något om jämförbarheten. I kravsammanhang kan man dock specificera krav som stödjer en sannolik jämförbarhet med andra mängder, t ex att samma mätskalor som för tidigare insamlade data används för ny datainsamling, att veder tagna nomenklaturer används, och givetvis att begreppsdefinitioner samordnas. Jämförbarheten är ju grunden för det mesta DA/IA-arbete.

Ett exempel på problem med jämförbarhet är när en forskare tar fram nya mätskalor för ett område där det finns tidigare statistiska data från andra undersökningar. Projektets ambition är hög och de tidigare använda mätskalorna förfinas i det nya projektet. När sedan en statistisk studie över en längre tid ska genomföras är resultaten från de olika undersökningarna inte jämförbara (sammandrag av exempel i Brammer 93).

Tillgänglighet är en systemegenskap som även kan komma att tas upp som en faktor i användarens bedömning av informationskvaliteten. Här kan man avse olika saker. Teknisk tillgänglighet har med det tekniska systemets egenskaper att göra. Logisk tillgänglighet – begriplighet i presentationen har också med det tekniska systemets utformning att göra, men även med användarens förutsättningar. Behörighet och ekonomisk tillgänglighet har med systemets administrativa aspekter att göra.

Kvalitetskrav vid nyutveckling och nyanvändning

Fastställande av krav på data- och informationskvalitet kan ske vid olika tillfällen.

Vid nyutveckling av data- informationsstrukturer. När data och register ska definieras och data börja samlas in. Detta aktualiseras som regel i samband med att nya system ska byggas upp, men det kan även vara ett led i en generell översyn av informationsresurserna i en verksamhet där en informations strategi ska fastställas.

I enlighet med IA-tankarna ska nyutvecklingen göras med tanke på ett bredare spektrum av användningsområden. Nyutvecklingen har initierats i ett syfte, bredare eller smalare. När data ändå ska samlas in och man är i kontakt med en viss uppgiftslämnare, frågar man sig vilka andra användningar som kan vara aktuella och vilka ytterligare data som eventuellt kan vara lämpliga att samla in. Men man är givetvis försiktig med sådan insamling "för säkerhets skull". Att samla in mer data än nödvändigt innebär ökade kostnader, också för informationsgivaren. I samband med nyutveckling gör man också undersökningar av vilka befintliga data som kan nyanvändas.

Vid nyanvändning (eller återanvändning). Existerande data används i ett nytt syfte eller i ett nytt system.

I en nyanvändningssituation som i och med att IA etableras kommer att bli allt vanligare, prövas lämpligheten för den nya användningen. Detta bör ske genom att kvalitetskraven specificeras först och att man därefter prövar om befintliga data kan användas. Eventuellt kan förändringar behöva göras i befintliga system och datamodeller.

Vid en generell kvalitetssatsning, där även informationskvalitet bör vara en del. Kvalitetskrav på och i verksamheten härleds vidare till krav på data- och informationskvalitet. Informationskvalitet specificeras för de data som krävs för själva kvalitetsarbetet.

En satsning där en stor del av verksamhetens data och begrepp går igenom. Det kan t ex röra sig om ett stadsplane- eller generalkartearbete. Kvalitetskrav och -specifikationer förs in som en del i arbetet.

Knyt ihop krav på informationskvalitet med verksamhetens krav

Kraven på informationskvalitet grundas på verksamhetens kvalitetskrav och produktivitetsambitioner. Krav på hög aktualitet hos data kan t ex grundas i ett behov av att kunna följa upp och styra verksamheten snabbt i realtid, eller på ett servicekrav mot kunder som innebär korta ledtider från order till leverans.

Dataanvändning för olika ändamål ställer olika krav. I fall där enskilda kan drabbas genom brister i rättssäkerheten är kraven helt andra än vid massutskick. Kvalitetskrav och -mått för verksamheten, samt produktivitetsmått och övriga resultatmått och -mål bör alltså finnas tillgängliga när kraven på informationskvalitet ska fastställas. Man bör även ha gjort bedömningar av vilka faktorer som är kritiska framgångsfaktorer i verksamheten.

Förhoppningsvis finns allt detta beslutat och dokumenterat. Annars kan identifiering av kvalitetskraven och bedömning av styrkan göras i samband med en första målmodellering av den aktuella verksamheten. I idealfallet bör även mätning och uppföljning av verksamhetens resultatmått inklusive kvalitetsmått ha påbörjats eller planerats. På så sätt ser man till att det finns data från tiden före den påbörjade användningen av de data som man ska specificera kvalitet för.

Vid fortsatt målmodellering tar man fram samband mellan verksamhetens kvalitets- och produktivitetskrav och krav på informationskvalitet. I begreppsmodellering, datamodellering och eventuellt i flödesmodelleringen¹⁸ identifieras vilka kvalitetsfaktorer som är aktuella för olika slags data och information.

De olika kvalitetsfaktorerna för information, noggrannhet, detaljeringsgrad, aktualitet, objektiva fel (feltolerans) tillgänglighet m fl, lyfts fram som en checklista för verksamhetskraven och en prövning sker. Prövningen bör alltså ske som ett extra arbetssteg i modelleringen och med verksamhets-kunniga personer närvarande. Ansvaret för att uppfylla kraven på data och information fördelas sedan på systemdelar och olika aktörer. Krav på t ex logisk tillgänglighet som påverkar informationsbildning hos användaren är beroende av systemets egenskaper. Noggrannheten hos insamlade data beror på hur själva mätningen och datainsamlingen går till.

En särskilt utsedd person bör dokumentera hur verksamhetens krav uppfylls genom specificerade kvalitetsnivåer på data och information.

Krav från olika intressenter och användningsfall

En kartläggning av intressenter bör göras. Denna kartläggning bör omfatta såväl intressenterna i den aktuella nyutvecklingen som tänkbara eller sannolika intressenter för senare nyanvändning.

¹⁸ En flödesmodell av processtyp kan t ex användas för att studera ledtider och hur de kan förändras. Ledtiderna påverkar vilken högsta nivå på aktualitet hos erhållna data som är möjlig att få.

För de olika intressenterna kan en genomgång göras av:

- Kända avsedda syften/användningsfall
- Tänkbara syften/användningsfall
- Olämpliga, inte tillåtna eller inte tänkbara syften/användningsfall

Identifikation av *primär*- och *sekundär*användare innebär att man gör klart för sig vilka användare eller intressenter som i första hand ska få sina krav uppfyllda och vilka som får hålla till godo med den kvalitet som kan ges. Sekundär användarnas krav uppfylls bara så länge de inte står i strid med primär användarnas krav eller om det inte innebär någon extra kostnad. Om man anser sig ha identifierat flera primär användare som har motstridiga krav får man pröva följande:

- Är den användare som ställer de mest kostsamma kraven beredd att helt stå för den tillkommande kostnaden?
- Kan nedbrytning i basdata och aggregering och behandling för olika syften vara en lösning (även här kan tillkommande kostnader uppstå)
- Behöver förhandlingar ske eller kompromisser göras i begrepp, definitioner och namngivning?

Terminologiruta

”**primär användare** användare för vars syfte informationen insamlas. I princip skall kravspecifikationen uppfylla primär användarens behov i första hand.”

”**sekundär användare** användare som använder information som tagits fram i annat syfte.

Användning kan vara möjlig därför att kravspecifikationen ser likadan ut trots att syftet är ett annat eller därför att sekundär användaren kan acceptera skillnader.”

ITS 94

Så många intressenter som möjligt bör medverka i det fortsatta kravarbetet direkt eller som remissinstans. Juridisk expertis kan även behöva medverka exempelvis i de fall då datakvaliteten har förvaltningsrättsliga aspekter.

Basregister eller inte?

Man bör klargöra om det register eller de datamängder som diskuteras har karaktär av basregister eller basdata: Vill man ha en gemensam informationsresurs och kopplingspunkt för flera organisatoriska enheter eller vill man i första hand använda data specifikt för en viss del av verksamheten? När data har stor grad av gemensamhet krävs en inriktning på större grad av nedbrytning och beredskap för flexibilitet.

Förutsägbarhet och stabilitet i kraven

En bedömning av förutsägbarhet och stabilitet kan behöva göras. Här får man se både bakåt och framåt och på så sätt bedöma om det behövs en säkerhetsmarginal för framtida krav. Erfarenheter av den verksamhet som berörs kan visa om det har varit vanligt med svängningar som kan motivera ändrade krav. Kända eller förväntade förändringar i framtiden kan rapporteras av de olika intressenterna.

Avvägning mellan olika kvalitetsfaktorer (pay-off)

Det kan vara svårt att med en rimlig resursinsats uppnå maximal nivå på alla kvalitetsfaktorer samtidigt. Det bör beaktas vid analysen så att man från början gör en prioritering. Verksamhetens kritiska framgångsfaktorer ger svar på vad som ska prioriteras.

Exempel på faktorer som man kan behöva avväga mellan är:

- Detaljeringsgrad eller noggrannhet kan stå i motsats till aktualitet (snabbhet). En grov resultatredovisning för en verksamhet kan tas fram på kort tid. Det kan räcka för att man ska veta hur det går och få tidiga varningssignaler om något ovanligt håller på att hända. Om man vill ha en detaljerad, noggrann och välkontrollerad redovisning så tar det tid.
- Jämförbarhet kan stå i motsats till anpassning för ett specifikt behov. Specifik anpassning av data till ett aktuellt behov kan medföra svårigheter att dra slutsatser från jämförelser med tidigare insamlade data.
- Detaljerad kan stå i motsats till noggrann. Mätning på en detaljerad skala ger större risk för fel.
- Specialiserade data kan stå i motsats till basdata. Specialiserade aggregerade data kan användas direkt i ett visst syfte. De tar lite resurser och plats att registrera och lagra. Basdata behöver behandlas innan de används. Att registrera och lagra basdata kräver resurser.

Data från tidigare resultatuppföljningar

Som en grund för kravställning tar man fram data som skapats i berörda verksamheter vid mätning och uppföljning av verksamheternas produktivitet och kvalitet. Undersökningar görs av vilka fel och brister som kan härledas till brister i informationskvaliteten. Problemanalyser som genomförts är också ett material som kan analyseras ur ett informationskvalitetsperspektiv.

Vad är rätt och vad är fel?

Kvalitetsbedömningen är som vi redan konstaterat starkt knuten till syftet, användarens uppfattning och effekter av informationsanvändning inom en viss process. Därför är det viktigt att klargöra vad som betraktas som rätt och fel, som bättre eller sämre kvalitetsnivå för varje syfte, part och användning. Är t ex en mindre felstavning i en adress ett fel? Hur stort mätfel är tillåtet inom ramen för "helt rätt"? Detta bör dokumenteras utförligt och behandlas i en kvalitetsgenomgång (review). Dokumentationen ger sedan hjälp med att svara på frågan om relevans hos data, dvs om de är ändamålsenliga för den aktuella användningen – givetvis förutsatt att förekomsterna stämmer med sin definition (validiteten).

Noll fel kan bara nås om man vet vad som är rätt och vad som är fel.

Risk- och effektanalys

Risker för fel, effekter och sannolikheter kartläggs och dokumenteras. Detta ger en grund för det fortsatta kravarbetet.

Felen kan (som vi redan sett) delas in i två kategorier:

- a: En avvägning kan göras mellan t ex en viss mängd fel eller brister i ett register och kostnaden för att undvika fel. Felfaktorerna identifieras, kvantifieras och värderas.

Exempel:

- Kostnader för effekter av kvalitetsbrister
- Kostnader för att åtgärda brister
- Kostnader för att upprätthålla en viss kvalitetsnivå eller kvalitetsklass

Man kan då räkna ut en optimal nivå på kvaliteten.

- b: Även enstaka fel är oacceptabla och noll fel måste uppnås. Denna kategori anknyter till det tidigare nämnda grundantagandet att en höjning av kvaliteten fram till uppnådda noll fel alltid är lönsam.

Båda synsätten har naturligtvis både svaga och starka punkter. Om man utgår från vad man specificerat som rätt och fel, gör man kalkyler för att fastställa feltoleranser eller också fastställer man noll fel som norm.

Vid varje bedömning bör man ta ställning till och göra bekant vilket synsätt som ska tillämpas. För olika delområden kan olika val bli aktuella.

En poäng med denna prövning är att man från början utgår ifrån möjliga fel och brister, och ser efter vad dessa kan leda till, i stället för att blunda inför risken och hålla tummarna. Indelningen kan göras i distinkta kvalitetsnivåer för olika faktorer, med t ex noll fel, delvis användbar information och helt oacceptabel nivå.

Nyttoanalys

Nyttoanalysen ger en grund för den positiva delen av den påföljande värderingen. Analysen kan göras med utgångspunkt i befintliga mätdata. Om sådana data inte finns utan det gäller att ta fram idéer om förväntad nytta så bör det naturligtvis dokumenteras så att en uppföljning kan göras.

Finansierings- och marknadsanalys

En undersökning kan behöva göras av hur kostnaderna för kvalitetsnivåer ska finansieras. En uppdelning kan t ex göras i:

- informationsgivande som kan avgiftsbeläggas, och där den enskilde betalande styr kvalitetsklasser samt bekostar eventuell anpassning
- informationsgivning som ska finansieras med anslag eller schablonavgifter och följa en viss beslutad kvalitetsnorm

Effekter av olika finansieringsformer med interndebitering osv studeras. Man bör t ex tänka över om avgifter kan få en negativ inverkan på den totala kvaliteten genom suboptimering.

Flexibilitetsanalys

Flexibilitetsanalysen handlar om vilka variationer i kvaliteten som är önskvärda respektive möjliga att åstadkomma – allt efter behov. Kan kunder eller användare vid olika tillfällen ändra nivå på kvalitet och kostnader? Kan en insats av extra resurser i det befintliga systemet ändra dess egenskaper så att t ex ledtiderna kan kortas eller antalet fel minskas?

Man måste också studera vad som styr eller begränsar den högsta möjliga kvalitetsnivån. Det kan t ex vara detaljeringsgraden i den skala som används eller noggrannhet i mätningen.

Exempel på andra frågor är:

- vilka möjligheter finns att minska kostnader om kvalitetskraven sänks?
- vilka specialanpassningar – ”förädlingar” – är möjliga att göra utifrån grundmaterialet?
- vilken flexibilitet, frihet från låsningar, kan åstadkommas utan stora extra kostnader?

Uppföljning av kravarbetet

Åtgärder som vidtas för att påverka informationskvaliteten bör i fortsättningen följas upp genom mätning i verksamhetens resultatmätt och nyckeltal.

Resultat från arbetet med att fastställa kvalitetskrav på data och information

Resultatet från kravanalysen sammanfattas och struktureras. Exempel på punkter som kan ingå är följande:

- Dokumentation i listform av hur respektive verksamhetskrav uppfylls genom specificerade kvalitetsnivåer på data och information. Gärna också vilka systemfunktioner som ansvarar för uppfyllande av respektive krav.
- Kvalitetskrav på grundmaterial eller basdata.
Härledda krav
 - Insamlings- och skapandeprocess
 - Kontroller
 - Indata som skaffas från annan källa
- Kvalitetskrav på bearbetade data.
Krav på systemets prestanda, kontroller, säkerhet m m. Gäller även manuella delar av systemet, organisation kring systemet osv.
- Krav på flexibilitet.
 - inom ett befintligt system med kort varselinklusive
 - förändringar som sker inom en viss tid (inklusive plan för förändring)
- Mätmetoder och mätskalor för kontroll av kravuppfyllnad.
- Testplan för data- och informationskvalitet vid systemtest.
- Redovisning av hur fel ska kunna spåras¹⁹.
- Uppföljningsplan för kontroll och mätning av kravuppfyllnad och effekter på produktivitet och kvalitet i verksamheten.
- Redovisning av beslut och beslutsmotivationer för beslut som hanterar problematiska frågor uppkomna under arbetets gång. Det är särskilt viktigt när man avvägt mellan olika intressen eller när man tagit beslut om att inte uppfylla de ursprungligen önskade kravnivåerna.

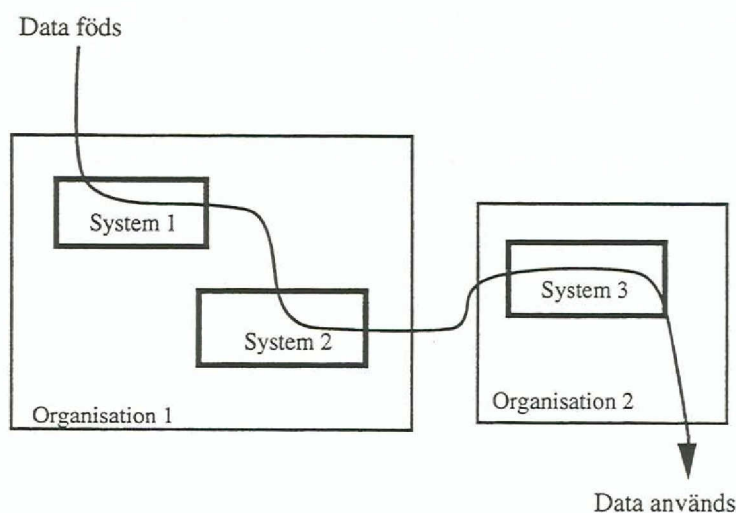
¹⁹ Ett revisionsmässigt kvalitetskrav är att kunna spåra hur data behandlats och till vilka system/register data överförs. Detta gäller historiskt redovisningsdata men också i princip alla uppgifter som en revisor kan vilja kontrollera. I och med årlig redovisning och resultatredovisning berörs många data av detta krav.

6 Arbeta med informationsprocesser, ansvar och intressen

6.1 Långa och sammanflätade datakedjor ger kvalitetsproblem

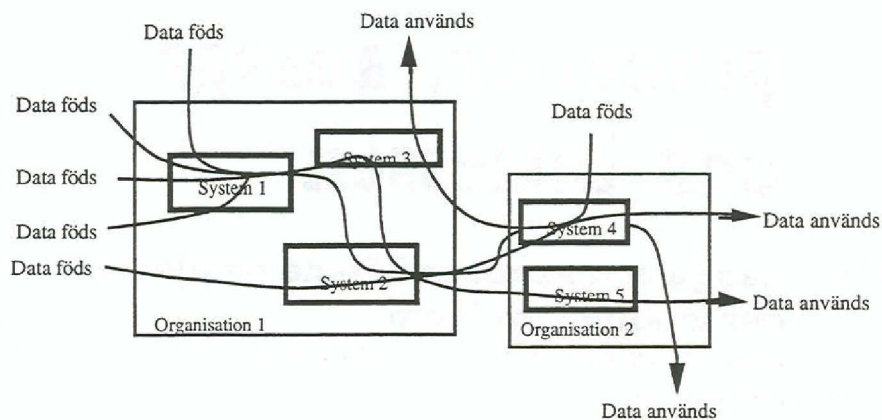
Det är vanligt att data passerar organisatoriska gränser under sin väg från insamling till användning.

Data föds, passerar genom ett antal funktionella system med eller utan behandling eller förädling och används till slut (se ill). De olika systemen har olika ansvariga. Det har även de olika organisationerna. Ansvaret för helheten kan därför vara oklart och splittrat. Det kan vara oklart vem eller vilka i kedjan som har ansvar för definitioner, för att data som överlämnas till en annan funktion är korrekta enligt sin definition osv. Även om det inte finns något beslut att göra så, har man skapat ett nytt system – ett *interorganisatoriskt informationssystem* – som sträcker sig över ett antal organisatoriska gränser. Det tidigare nämnda adresssystemet (SAIS) är ett exempel.



Figur 5 – Datakedja eller informationslinje från födelse till användning.

När data från olika källor sätts samman för olika användningar blir det hela mer komplicerat. Situationen kan t ex leda till att datamängder som inte är jämförbara ställs samman och jämförs.



Figur 6

Ansatser av typen "information warehouse" och liknande, innebär att data från olika håll är tillgängliga gemensamt för en användare. Det finns då dubbla risker, dels att data används i ett nytt syfte som de kanske inte passar för, dels att data som inte är jämförbara ställs samman och jämförs. Den som ansvarar för ett "information warehouse" kan alltså inte nöja sig med att vidareförmedla data, utan måste också ta ansvar för att datas betydelse, kvalitet och lämplighet för olika användningsområden står klara för mottagarna.

Problembilden skulle kunna tolkas så att det vore bäst att alltid hämta data från ursprungskällan. Det kan dock vara svårt eller opraktiskt. Lösningar med mellanhänder kan behövas.

Ansvaret för dataflödet från "ax och gräs till ostmacka i magen" är dock ofta splittrat.

6.2 1:1:1:1-principen

Ett synsätt som lanserats för att tackla detta problem, är "1:1:1:1-principen" (Andersson och Nilsson 93). Andersson och Nilsson beskriver problemet som "complexity in inter-organisational information systems" (IOIS). Som lösning föreslår man en ny syn på vad som ska betraktas som ett informationssystem. Det innebär att man använder ordet i en speciell betydelse. Storleken på det som man betraktar som ett system måste reduceras för att det hela ska bli överblickbart. Systemet ska bara ha ett syfte och detta syfte ska vara klart uttryckt och helst tidsbegränsat. Om det finns flera syften så ska man definiera flera system. 1:1:1:1-principen innebär alltså att *ett* informationssystem ska höra till *en* aktivitet för *en* part och ha *ett* specifikt syfte inom denna aktivitet och det man menar med systemet är själva informationsflödet.

Det kanske dock kan sättas ett frågetecken för lämpligheten i att försöka ge det allmänt använda uttrycket "informationssystem"²⁰ en ny betydelse. Ett mer eller mindre nybildat ord kanske skulle kunna användas, t ex informationsprocess, datakedja eller informationslinje. Eller kanske "1:1:1:1-system" för att anknyta till Andersson och Nilssons behandling av ämnet.

Tanken är dock värdefull. Framför allt är den användbar som utgångspunkt för analys och diskussion när ett flertal parter är inblandade. Resultatet är inte givet som en uppstyckning i ett eller flera "1:1:1:1-system" utan snarare som medvetenhet och hänsynstagande till de problem som komplexa system innebär.

En annan fråga är om 1:1:1:1-principen kan förenas med IA- och DA-ansatser. Att identifiera 1:1:1:1-system bör kunna göras i en analysfas och som en förening av DA-arbete och process management. En tanke skulle kunna vara att utse ansvariga för de olika informationsprocesserna eller informationslinjerna. Inom området process management studeras företagets olika produktionsprocesser. I ett tillverkande företag går processen för att tillverka en viss produkt genom ett antal olika funktioner: materialinköp, verkstad, målning, lager osv. Man försöker öka effektiviteten genom att kartlägga processer, reparera trasiga processer och utse processägare som är ansvariga för sina respektive processer i sin helhet.

²⁰ Termen informationssystem används säkert med varierande innebörd beroende på vem som talar. Den betydelse som läggs i ordet innefattar dock ofta troligen programvara m m och inte bara själva dataflödet som Nilsson och Andersson fokuserar på.

Andersson och Nilsson ser inte ett sådant övergripande ansvar för hela processer som en lösning. Man betonar i stället behovet av samverkan mellan de inblandade aktörerna. Dessa måste enligt Andersson och Nilsson:

- definiera syftet med det interorganisatoriska informationssystemet klart och entydigt
- skaffa sig överblick över och inblick i det interorganisatoriska informationssystemet
- bevaka förändringar i omgivningen effektivt för att kunna upptäcka betydelsefulla förändringar i god tid
- utveckla gemensamma principer, strategier och avtal
- definiera ansvarsförhållanden mellan de olika aktörerna och för delarna i det interorganisatoriska informationssystemet
- utveckla effektiva utrednings- och samarbetsprocesser både i utveckling och administration av interorganisatoriska informationssystem

6.3 Svårigheter att förena intressen

Parter i en informationsprocess kan ha intressen som inte är helt förenliga eller till och med motstridiga. Motsättningar mellan användare av samma data är ett exempel.

Om data ska samlas in eller system ska byggas för att tillgodose flera olika användarintressen, kan det vara svårt att tillfredsställa alla samtidigt. Parterna kan ha olika uppfattningar om terminologi, mätmetoder, kvalitetsnivåer osv. Man kan bli oense om vem som ska bära kostnader för kvalitetsnivåer. Där ett val måste göras mellan olika faktorer, kan önskemålen om prioritering gå isär. Om projektet drivs på fel sätt och med orealistiska förhoppningar om att kunna förena många intressen, kan resultatet bli en kompromiss där ingen part är nöjd.

Ett exempel är informationssystemet för arbetsskador, ISA, som fick en så bred målformulering att ingen användare var nöjd med den tillgängliga informationen (RRV 1990).

Ett annat exempel på svårighet att förena intressen är när det finns en motsättning mellan uppgiftslämnare och -mottagare.

Motsättningen mellan uppgiftslämnare och uppgiftsmottagare kan leda till att uppgifter inte alls kommer in eller att uppgifter lämnas men är ofullständiga och mer eller mindre felaktiga.

Två typer av motsättningar kan urskiljas. I det första fallet är det så att det inte finns intresse av att lämna uppgifter eller av att kontrollera eller anpassa uppgifterna så att de får den kvalitet som mottagaren behöver. Orsaken är att uppgiftslämnaren har besväret med att lämna uppgiften, men inte får ut något av det hela. Nyttan och uppoffring finns på olika ställen.

I den andra typen av motsättning finns mer direkt motstridiga intressen eller syften mellan parterna. Det finns ett intresse hos uppgiftslämnaren att undanhålla uppgifter eller lämna falska uppgifter.

Personer kan t ex av skatteskal ha intresse av att vara folkbokförda på en annan adress än den där de verkligen bor och egentligen vill ha sin post till. Eftersom anmälan av folkbokföringsadress och postadress hanterades gemensamt i det statliga adressinformationssystemet (beskrivet i RRV 92) fanns risken att vissa personer fick fel postadress i ett antal register.

De politiska ungdomsförbunden lämnar felaktiga uppgifter om medlemsantal av ekonomiska skäl och konkurrensskal (Dagens Eko 16 & 17 januari 1994). Detta innebär att sammanställningar av medlemsantal i ungdomsförbunden inte kan användas som en säker indikator på politiskt intresse hos ungdomar osv.

I detta sammanhang finns ytterligare en faktor att beakta. Uppgiftslämnaren kommer att rätta sitt agerande efter *sin uppfattning om* insamlarens syften. Om t ex en myndighet samlar in uppgifter i ett i och för sig oskyldigt syfte, kan uppgiftslämnaren oroa sig för att uppgifterna lämnas vidare till någon annan myndighet, oavsett om det kan ske enligt gällande regelverk eller inte. Om syftet, och de regler som gäller för hantering av uppgifterna, redovisas oklart, kan osäkerheten hos uppgiftslämnaren också bli stor.

Ett tredje exempel är när en part i kedjan är beroende av en tidigare part, som kanske både samlar in uppgifter och bearbetar dem. De brister som gör att data inte är lämpliga för den mottagande partens syfte kan upptäckas vid användningen eller vara kända i förväg. Detta kan t ex röra sig om att data har fel form, fel detaljeringsgrad eller att de kommer in för sent, dvs inte är tillräckligt aktuella. Användaren kan dock inte göra något åt saken, eftersom det inte står i hans eller hennes makt att påverka den som levererar data. En verksamhetsdel får då bära kostnader för bristande kvalitet i en annan. Det kan röra sig om avdelningar vid ett stort företag eller olika myndigheter.

Ett exempel är när en myndighet i sitt arbete använder statistiska data som en annan myndighet samlar in. När arbetstoppar uppstår hos den insamlande myndigheten så försenas leveransen av data till den mottagande av prioriteringsskal. Den användande myndigheten måste då själv samla in samma data för att kunna leverera uppgifter till sina uppdragsgivare i tid.

Försök förena intressen

Med hjälp av kreativa upplägg kan man försöka förena olika parter intressen. Det kanske inte alltid är möjligt att göra det men man bör ändå prova, eller försöka skapa förutsättningar för det.

Några generella anvisningar för hur det ska gå till kan inte ges här, men ett exempel kan belysa saken:

I trafikregistret är man angelägen om att ha så aktuella ägaruppgifter som möjligt för de registrerade fordonen. I det syftet har man lyckats förena sina intressen med bilhandlarna. Bilhandlarna registrerar ägarbyten per terminal direkt efter att de sålt bilen. Bilhandlarens intresse av att göra det är bland annat att ge köparen en god service, denne behöver inte bekymra sig om att registrera sig som ägare för bilen. Dessutom upphör bilhandlaren betalningsansvar för försäkringen så fort en omregistrering skett.

I vissa fall kan förändrade regler skapa den struktur som behövs för att förena intressen. Fundera t ex på hur denna regel skulle påverka agerandet:

”Fabrikens vattenintag måste enligt lag ligga nedströms dess eget vattenutsläpp så att den blir den första som drabbas om utsläppen inte renas” (deBono 84)²¹.

Om det inte går att förena olika intressen måste man vidta andra åtgärder. I vissa fall kan det vara nödvändigt att separera informationssystem för olika intressen i syfte att upprätthålla önskade kvalitetsnivåer. I andra fall kan lösningen vara avtal, överenskommelser om ersättningar, nedbrytning i bas data, uppdelning i olika organisationsdelar där vissa renodlar rollen som ”datanav²²” osv.

²¹ Läsaren uppmanas, att som övningsuppgift, tänka ut en motsvarande regel för informationssystem.

²² Med ”datanav” menas här att man är en växlingspunkt där en mängd dataflöden korsar varandra.

Klargör ansvar och dess fördelning

Vem ansvarar för att data håller en viss kvalitet? Det kan vara svårt att t ex kräva ekonomiskt ansvar av den som levererar data.

Det är givetvis möjligt att se på ansvarsfrågan på olika sätt. Ett sätt är att se på den avtalsmässiga och ekonomiska sidan av saken. Där kan det bli fråga om skadestånd eller avdrag på ersättningen om en kontraktsbunden part inte uppfyller sitt ansvar. Ett annat troligen mera fruktbart synsätt är att klargöra hur ansvar och befogenheter ska fördelas inom organisationen. Där är det viktigt att få verksamhetsansvariga att uppmärksamma sitt ansvar för de data, begrepp osv som hör ihop med verksamheten.

I ITS 6 skiljer man på tre olika slags ansvar:

"Dataleverantörsansvar"²³, ansvar för att data som levereras till systemet är riktiga". (ITS 94)

I en kommentar sägs att "Individ med detta ansvar är strikt den ende som kan vara ansvarig för att data som lämnas för inmatning är riktiga enligt specifikation. Denne behöver inte ha någon annan koppling till verksamheten eller själv använda uppgifterna" (ibid).

Man kan tolka formuleringen så att det är leverantören av data som har störst *möjlighet* att ta ansvar för kvaliteten hos de data som levereras. Som vi sett kan möjligheterna att få någon extern part att *vilja* ta detta ansvar variera. Klargörande av intentioner, i kombination med goda kontakter med dessa parter, är troligen en bättre väg att gå än att försöka göra strikt juridiska bindningar. Poängen med formuleringen är att mottagaren av data kan ta sitt ansvar genom kontroller, genom att kräva och studera kvalitetsdeklarationer osv. Mottagaren har däremot betydligt mindre möjlighet att kontrollera och mäta kvaliteten än leverantören.

Det är speciellt viktigt att se till att dataleverantören ansvarar för de kvalitetsdeklarationer som lämnas. Genom att t ex föreslå dataleverantörerna att sådana deklarationer upprättas ger man dem tillfälle att uppfylla en grundläggande nivå av ansvar.

"Definitionsansvar, ansvar för att definiera (och ändra) de objektsbeskrivningar som ingår i ett register". (ITS 94)

Detta ansvar bör i grunden ligga hos själva verksamheten. I praktiken kan det utövas av en DA- eller IA- funktion.

Detta är dock ett administrativt ansvar. Någon absolut makt över språkbruket i en verksamhet ger det inte.

²³ Detta kallades tidigare "indataansvar", en term som dock skulle kunna misstolkas som "den som har ansvar för att ta hand om de data som kommer in"

”Registeransvar, ansvar med rätt att uppdatera eller ändra data inklusive arkivansvar.

Registeransvar kan i praktiken finnas hos systemägaren eller hos uppgiftsinsamlaren. Vem som har rätt och därmed ansvar att läsa, ändra, ta bort etc måste alltid fastställas. Jämför lagtermen **registeransvarig** med annan innebörd.” (ITS 94)

Registeransvaret bör troligen även innefatta ansvar att undersöka om det finns olika slags validitetsbrister i de register man har ansvar för, om möjligt åtgärda dessa och annars informera användarna om dem i kvalitetsdeklarationen.

Vidare definieras ett vidare begrepp ”informationsägare” enligt följande:

”informationsägare, individ med ansvar för informationen i en verksamhets funktion. Ansvaret kan innefatta såväl **definitionsansvar**, **registeransvar** som **dataleveranstörs ansvar**.” (ITS 94)

Användarens – beslutsfattarens – ansvar och befogenheter

Ett annat ansvar som kan vara värt att beröra är det ansvar som en beslutsfattare har för att kontrollera de data man använder, t ex att de har en trovärdig kvalitetsdeklaration.

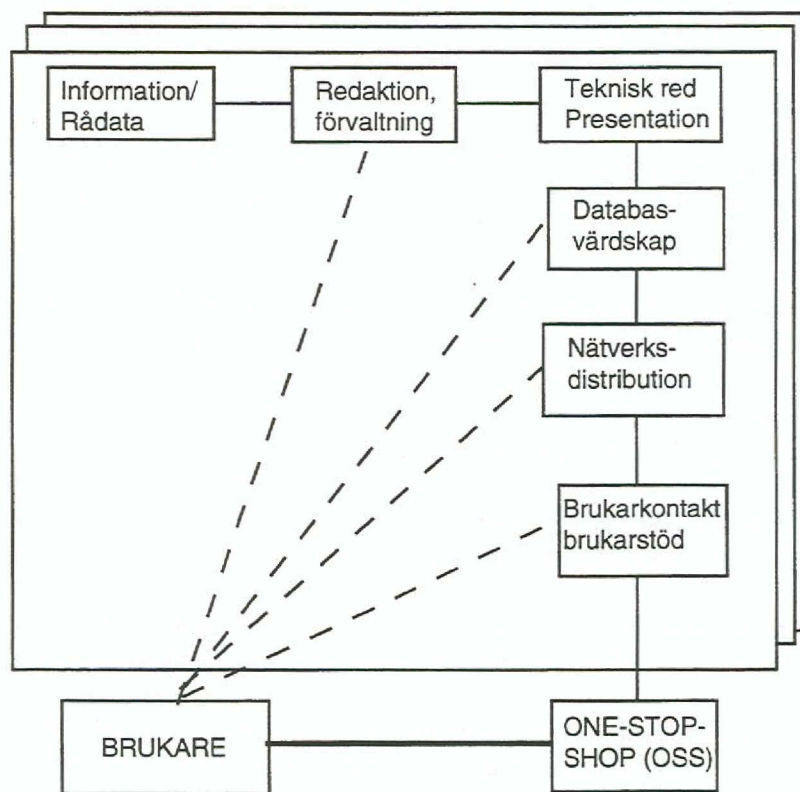
I en reguljär rutinartad verksamhet kan man troligen inte kräva att den enskilde beslutsfattaren ska kontrollera data vid varje enskilt beslut. Här bör något slags system för kvalitetsstyrning byggas upp. Större krav ställs när informationskällor utnyttjas för tillfälligt bruk. Då måste användaren – beslutsfattaren – själv vara aktiv i sin aktörsroll och undersöka kvalitetsdeklarationer osv. Man måste utgå från att data är fel tills motsatsen är bevisad. Ansvaret innebär dock också att användaren eller aktören måste ha befogenheter; både att ställa krav på data i sig, kvalitetsdeklarationer osv och få dem uppfyllda och att avstå från beslut eller annat handlande när han bedömer det nödvändigt.

Beslutsfattaren måste i särskilt viktiga beslut ha möjlighet att göra extra kontroller mot säkrare källor och även få återkoppling på effekterna av sin informationsanvändning. För att användaren ska kunna ta sitt ansvar, krävs också utbildning i frågor kring informationskvalitet. Det kan även finnas skäl till att beslutsfattare dokumenterar vilken information de använt och med vilken deklaration, vilka extra kontroller de gjort, om beslut av något skäl har framtvingsats trots tveksamheter på grund av osäker datakvalitet osv. Beslutsfattaren kan då i efterhand redovisa hur han eller hon har axlat sitt ansvar.

Att enbart ge användarna information om kvalitetsbrister *utan* att ge dem kunskaper och befogenheter att agera efter detta, försvårar dock deras situation. Om användaren är medveten om att kvaliteten är osäker och ändå inte har något annat val än att fatta beslut finns det risk till s k inlär hjälplöshet hos användarna. Användaren blir då inte alls någon *aktör* utan en maktlös *operatör* i systemet. Detta blir på sikt förödande för kvaliteten i verksamheten.

Hur kan roller och uppgifter i kedjan utformas och renodlas?

I en delrapport i projektet "Nasjonal infrastruktur for EDB" i Norge (Statskonsult 91) skisseras en uppdelning på olika roller i förmedlingskedjan:



Figur 7 – Bild av förmedlingskedjan, Statskonsult, Norge.

Huvudaktörerna i kedjan identifieras som:

Informationsleverantör – den som producerar de grunddata som distribueras i en förmedlingskedja

Tjänsteleverantör – den som utifrån eget eller andras data-insamlande ger informationstjänster till andra med eller utan avgift

Distributör – den som står för förmedlingen av data mellan tjänsteleverantören och brukaren

Brukare – en person som använder en eller flera förmedlingskanaler för att nå ett tjänsteutbud från en leverantör/distributör

Statskonsult har även tagit fram standardavtal som reglerar förhållandena mellan de olika parterna.

Avtal

RRV rekommenderar avtal (RRV 90) som ett sätt att lösa vissa kvalitetsproblem, framför allt sådana som har med ansvar för informationskvaliteten att göra.

I kommersiellt tillgängliga databaser²⁴ fränsäger sig vanligen databasproducenterna ansvaret för datakvaliteten och för effekterna av eventuella brister (Svensson och Östlund 1992). Försiktigheten hos databasproducenten är inte svår att förstå med tanke på de många möjligheterna till svårupptäckta fel och hur pass stora de negativa konsekvenserna i olyckliga fall kan bli. Det kan också vara svårt att avgöra vem som har ansvaret när flera parter är inblandade. En part samlar in data medan en annan part, t ex en databasvärd, förmedlar data till slutanvändaren.

Vilka möjligheter finns det då att teckna avtal där åtminstone datakvaliteten kan behandlas och vilka syften kan uppfyllas i avtalen?

Till att börja med kan man tänka sig en enklare nivå som kan sägas vara en skriftlig överenskommelse. I detta fall råder inte något affärsmässigt förhållande mellan parterna i fråga om dataöverföringen. Syftet är att skapa klarhet. I den skriftliga överenskommelsen behandlas de villkor som gäller för dataöverföringen, hur data kontrolleras och behandlas, vilken kvalitet som kan förväntas osv. Lämpligen medföljer en kvalitetsdeklaration från den part som lämnar data och en beskrivning av användning och syfte från den som tar emot data. I själva arbetet med att skriva överenskommelsen kan eventuella oklarheter kring begreppsdefinitioner och annat redas ut.

²⁴ Tillgängliga via uppkoppling eller i kopia via t ex CD-rom.

Förutsättningen för en sådan överenskommelse är att den givande parten har ett ansvar för att tillgodose den andres behov utan ersättning från den mottagande parten. Det kan röra sig om avdelningar inom samma företag eller olika delar av den offentliga sektorn. Man får dock vara beredd på att den part som tillhandahåller data i trängda lägen prioriterar sina egna huvuduppgifter²⁵ framför mottagarens behov. Förhoppningsvis hålls ändå de delar av överenskommelsen som stadgar att upplysningar ska ges om ändringar och avvikelser.

Nästa, högre nivå är mer formella avtal där syftet både är att skapa klarhet och att säkerställa vissa förhållanden och villkor. Avtalet är formellt bindande för parterna och den mottagande parten ska ge ekonomisk ersättning till den som lämnar data eller gör data tillgängliga. Ersättningen kan variera beroende på vilka tjänster dataleverantören tillhandahåller i form av kontroller, anpassning, bearbetningar osv. Rättigheter, skyldigheter och sanktioner regleras i avtalet.

Även här kan kvalitetsdeklarationen finnas med som bilaga.

När överenskommelser och avtal skrivs, bör man givetvis uppmärksamma de rättsliga krav som gäller för dataöverföring.

En tredje situation är när en dataleverantör erbjuder en standardiserad tjänst till en stor mängd användare mot en viss ersättning. I dessa fall har leverantören vanligen ett standardavtal som kan göra mottagarens ställning svag. Flera mottagare kan då behöva agera gemensamt i användargrupper eller -föreningar för att få acceptabla villkor – åtminstone i form av kvalitetsdeklarationer.

Överenskommelser och avtal om datautbyte kan klargöra vad som gäller och hjälpa till att undvika fallgropar. Själva arbetet med att ta fram dem ger en kvalitetssäkring, men det är svårt att få helt säkra garantier för en viss kvalitet. Det är troligen tillgången till upplysningar som ger den bästa garantin för att man ska kunna bedöma kvaliteten (deklarationer). Data-mottagarens eller användarens slutliga huvudansvar för att ha använt sig av kontrollerade och lämpliga data vid beslut kvarstår alltså.

Man kan dra en parallell till vad som rekommenderas i EDI-världen. Där sägs att goda affärsförbindelser och ömsesidigt förtroende är en grundförutsättning för att börja med EDI. Dessa goda förbindelser skapas inte genom avtal – avtalet är troligen snarare en bekräftelse på dem.

²⁵ I de fall där huvuduppgiften inte är att förse andra med data.

Den dataleverantör som tecknar ett avtal kan behöva överväga på vilka sätt han blir bunden i sitt agerande. Man bör undersöka påverkan av t ex möjligheterna att förändra system, lägga ner system och sluta samla in data man själv inte längre behöver. Åtagandena får inte ge bindningar, som hindrar viktiga förändringar i den egna verksamheten.

6.4 Kvalitetsstyrning

I kvalitetsstyrningen inom området data- och informationskvalitet kan man i huvudsak se två slingor:

Den första slingan innefattar sådana mätningar och åtgärder som görs inom ett existerande system utan att systemet behöver ändras. Slingan kan sägas utgöra en del av systemförvaltningen. Möjligheten att göra anpassningar inom denna slinga beror på hur systemet har byggts. Är systemet så pass flexibelt så att anpassningar kan göras, antingen genom att man så säga vrider på rattar i systemet eller genom att man arbetar annorlunda?

I detta innefattas att kontinuerligt mäta kvalitet och samla in eventuella klagomål. Även användarforum och liknande aktiviteter ingår. För en data-bastjänst med många olika användare kan man överväga att bilda en användarförening som bidrar till kvalitetsstyrningen. I slingan ingår också att informera användare och andra mottagare av data om den faktiska kvalitet som föreligger. En annan åtgärd kan vara utbildning i hur systemet fungerar och i informationskvalitetsfrågor. Slingan bör även omfatta indikatorer på kvalitet i den verksamhet som systemet eller systemen stöder.

Den andra slingan omfattar mer radikala ändringar i system, organisation och arbetssätt. Ändrade krav kan också vara en del. Antingen genom att den kvalitet som är möjlig utan ändringar i systemet accepteras eller att ändringar i kraven kräver större ändringar i systemet. Om de brister man finner vid en mätning inte orsakar problem i verksamheten, är kanske kvalitetskraven från början för högt ställda eller också behövs dessa data inte alls. Möjligheter till informationsrationalisering genom nya arbetssätt (eller "business process reengineering") bör också beaktas om man överväger en större ändring för att komma åt felkällor.

Skapa återkopplingar

En central del av arbetet med kvalitetsstyrning av informationsprocesser är att se till att det finns fungerande återkopplingar. Väl fungerande återkopplingar är en kritisk framgångsfaktor för alla system i en föränderlig omgivning och sådana är de flesta informationssystem. Brister i detta avseende leder till att fel och följderna av fel inte upptäcks. Därmed åtgärdas de inte heller. Effekterna av felet kan uppkomma utom synhåll för användaren eller kanske inte är så påtagliga att de uppmärksammas. Det är heller inte säkert att effekterna sätts i samband med brister i data- eller informationskvalitet.

Man kan kanske tro att de fel som upptäcks av användarna genast rapporteras men så är inte alltid fallet. Användarna upptäcker felet, men av olika skäl så rapporteras det inte. Exempel på troliga och tänkbara skäl kan vara:

- Användarna anser inte att felet är så stort att det bör anmälas.
- Användarna har inte tid eller ork att anmäla felet. De anser att de får nog med extra arbete ändå och hoppas att felet inte ska drabba dem igen.
- Användarna tror inte att felet kommer att åtgärdas även om det anmäls. De löser därför problemet på något annat sätt. En sådan attityd kan uppkomma genom användarnas tidigare erfarenheter av hur anmälda fel hanteras, t ex hur lång tid det brukar ta innan något görs och hur mycket de vidtagna åtgärderna förbättrar situationen.
- Användarna vet inte hur och till vem fel ska anmälas.
- Användarna tror att felet ligger hos dem själva. De tror att de borde förstå det som de inte förstår och de vill inte avslöja sin okunnighet.
- Användarna har låga förväntningar på systemet och hoppar högt av glädje när det fungerar väl. När något fel uppträder säger de rutinmässigt "fel på datan". Felen ses som något nödvändigt ont i det dagliga arbetet och man har aldrig tänkt på att det skulle gå att göra något åt dem.
- Dataleverantören eller den som är systemansvarig reagerar negativt på anmälningar om fel. De användare som anmäler fel ofta betraktas som kverulanter eller som för lata för att göra det extra arbete som krävs för att kompensera eller själv rätta felet när det uppträder.
- Dataleverantören eller den som är systemansvarig underlåter att kontakta användarna då och då för att få uppgifter om fel och förbättringsförslag, dvs användarforum saknas.

Nästa led handlar om återkoppling av felets effekter. Orsaker till det kan i första hand vara att användaren av data inte upptäcker felet och inte får någon återkoppling av de effekter som dataanvändningen orsakar i nästa led.

I nästa led kan flera punkter från listan ovan upprepas men ordet användare byts ut mot kund eller medborgare.

Exempel på vad som kan hända är:

- En kund får ett leveranstidsbesked som inte stämmer. Kunden klagar inte utan vänder sig nästa gång till en annan leverantör.
- En medborgare får ett felaktigt beslut från en myndighet, men kan inte avgöra om det är korrekt. Ett alternativt problem är att medborgaren förstår att beslutet är fel men vet inte hur man klagar.

Återkoppling från användarna till den som ansvarar för systemet och data-kvaliteten kan skapas genom punktinsatser som användarforum, kundundersökningar osv. Lämpligen bör man dock försöka bygga system för kontinuerlig återkoppling. Om mätningarna görs med vissa mellanrum, bör frekvensen anpassas till frekvensen av ändringar i systemet och i den omvärld i vilken systemet finns.

Slutligen behöver även användaren återkoppling på sin användning av information för att kunna anpassa sin användning och ställa välgrundade krav. Återkopplingen ska komma från mätningar och iakttagelser av det man avser att uppnå – mål och resultat.

7 Mät och informera om informationskvaliteten

Användarna behöver ha en korrekt kunskap om den data- och informationskvalitet som står till buds. För detta krävs att den faktiska kvaliteten mäts och att användarna känner till både *avsedd* och *faktisk* kvalitet.

7.1 Mätning av informationskvalitet

Mätning och uppföljning av data- och informationskvalitet kan göras på olika sätt, i olika syften och med olika ambitionsnivåer. En huvudindelning av syftena kan göras i *kvalitetsstyrning* och *information till användarna*. När syftet är kvalitetsstyrning, vill man ha mätdata som indikerar om kvalitetsmål och normer uppfylls och även sådana data som kan ge ledning för åtgärder, t ex indikationer på felorsaker. Uppmätta effekter av fel och brister kan också vara intressanta i sammanhanget. När syftet är "information till användarna," ska användarna få upplysningar som ger dem möjlighet att bedöma datas lämplighet för olika slags användning. De vill ha reda på precis vari felet består och kanske även vilka effekter det kan få. Användare kan vara av olika kategorier: handläggare som ska fatta beslut, informationsförmedlare, användande organisation (eller en företrädare). Mätning kan göras i ett av syftena eller i båda medan redovisningen av resultatet blir olika beroende på mottagarens syfte.

Mätning - kontroll mot verkligheten eller en säkrare källa

För att få en bild av vilka brister som förekommer kan man göra en statistisk mätning. Det finns olika sätt att göra detta. Ett sätt är att slumpmässigt välja ut en mindre mängd från det register som ska studeras och göra en kontroll mot en säkrare källa. Det kan röra sig om att kontrollera ett grundmaterial eller att göra en kontroll direkt mot verkligheten.

- Tidningsstatistik AB gör exempelvis årligen en sådan kontroll av ett antal adressregister. Man ber att få uppgifter från respektive registerhållares register (880 stycken per adressbank) och gör sedan kontroller per telefon. Antal rätt redovisas i procent. Uppgiften "bransch" var t ex i 1992 års undersökning rätt till 93% i BASUN och 83% i PAR.
- RRV gjorde en undersökning av fel i adresser och redovisar i "Fel data kostar!" (RRV 92) hur ett urval av försändelser under sortering gjordes på ett antal postkontor och hur man kontrollerade fel på olika sätt.

- Ett urval av klassificeringar kan undersökas med hjälp av en mer kvalificerad granskare. Det har man gjort i en studie av kvaliteten i Social styrelsens slutenvårdsregister (Socialstyrelsen 1992). Ett urval gjordes och en diagnosklassificering gjordes på nytt utifrån journaluppgifter. Därefter kunde en jämförelse göras mellan ursprunglig klassifikation som då antogs representera faktisk kvalitet i registret och i den mer kvalificerade bedömningen.

I "Rätt data?" (RRV 90) finns en genomgång av hur man kan mäta kvalitet med hjälp av små stickprov.

Iakttagelse av registeregenskaper

Med registeregenskaper menas sådant som man kan iaktta genom att under en viss tid göra observationer av data direkt i registret eller systemegenskaper som man annars känner till (t ex ledtider). Det kan t ex vara omsättningshastighet för vissa uppgifter i registret eller vanligen återkommande variationer.

Antag t ex att registret uppdateras satsvis en gång i månaden och att 120 objekt tillkommer i en jämn takt per år. Strax före en uppdatering kan man då i genomsnitt räkna med att cirka tio av de objekt som borde vara med i registret saknas. Om studier av registret visar att ändringarna inte är fördelade jämnt över året utan att majoriteten av nyregistreringar sker runt årsskiftet, ger man kanske sina användare rådet att vänta med sina uttag till efter januari uppdateringen.

Fullständighet, dvs om alla avsedda uppgifter om ett visst objekt finns med, kan iakttas direkt i registret.

Studier av registret i sig kan även innebära att andra uppenbara fel upptäcks, som t ex att det i ett adressregister finns dubletter där namnet på ett företag har stavats lite olika eller att helt felaktiga koder finns i urvalsfälten. Sådana fel kanske inte upptäcks så lätt eftersom det företag som får två eller tre exemplar av samma utskick inte hör av sig och klagar utan bara ställer sig lite undrande till hur det är ställt med ordningen hos den som har skickat ut försändelserna. De mottagare som på grund av fel i urvalskoder inte alls får någon försändelse hör troligen inte heller av sig.

Undersökning av vilken informationsbildning ett system ger hos handläggare och beslutsfattare

Detta kan sägas vara en undersökning som inriktar sig på att ta reda på om en viss given datakvalitet också omsätts i informationskvalitet. En kvalificerad tolkning av de data som finns i registret jämförs med den tolkning som görs vid ett normalt uttag, dvs ett sådant uttag som görs som ett led i de ordinarie arbetsuppgifterna.

En sådan undersökning gjordes hos trafikregistret. Den utformades som en serviceundersökning hos länsstyrelsen och gjordes med hjälp av ett externt företag. Personer från detta företag ringde och ställde frågor om de uppgifter som fanns i databasen. Sedan gjordes kontroller av om svaren som gavs var korrekta i förhållande till vad den kvalificerade bedömaren kunde läsa ut av databasen. Exempelvis angavs ägare 100%, avregistrering 85%, och besiktning 84% rätt.

Kvalitetsinformation till användare - kvalitetsdeklarationer

RRV (1990) och ITS (1994) rekommenderar kvalitetsdeklarationer för att informera om kvalitet. I rapporten ITS 6 ges även rekommendationer om vad som bör ingå i en deklARATION.

Kvalitetsdeklarationen har som syfte att upplysa användaren eller den användande organisationen om kvaliteten hos data på ett sådant sätt att lämpligheten för ett visst syfte kan bedömas. Alltefter mottagare och den situation där mottagaren tar del av deklARATIONEN, kan den ges mycket varierande utformning. En deklARATION som är avsedd att direkt upplysa en användare om kvalitet hos vissa data, får t ex en annan form än en som ska vara bilaga till ett avtal om överföring av data. Termen "deklARATION" för tanken till ett pappersdokument. Men också andra former för spridning kan mycket väl tänkas, t ex att deklARATIONsuppgifterna är åtkomliga via samma system som de data de handlar om. Det innebär att de skickas med data, unge får som bruksanvisningen till en produkt.

Databashanterare och datakatalog- eller IRDS-system är dock oftast inte byggda för att förmedla data om datakvalitet. Vissa uppgifter som har med relevansen att göra kan förmedlas i form av definitioner av hur vissa data härleds från andra osv ²⁶. Uppgifter som har med validitet att göra, statistik över felprocent eller registrets förändringsdynamik, har däremot ingen given plats.

Terminologiruta

KvalitetsdeklARATION beskrivning av data eller datamängd i syfte att användaren skall kunna göra en korrekt tolkning av sagda data.

ITS 94

²⁶ Härledning i AD/Cycle beskrivs i TRIAD-rapport K 15

Problem när användarna inte är informerade om informationskvaliteten

Bristande kunskaper hos användarna kan leda till att data eller information av bristande kvalitet används.

Det kan t ex röra sig om att data har objektiva fel som inte är kända för användaren och som är sådana att de inte upptäcks vid användningen. Data skulle ha varit lämpliga för syftet om det inte funnits objektiva fel.

Bristande aktualitet i t ex en adress kan vara en orsak till detta. För användaren gör det egentligen detsamma om felet beror på att adressen är gammal eller om adressen förändrats på ett felaktigt sätt i systemet.

Ett annat fall som i och för sig kunde ha förebyggts genom normalt DA/IA-arbete är att data som inte är lämpliga för det aktuella syftet används. En vanlig orsak är att data används i ett annat syfte än det ursprungliga och att lämpligheten för den nya användningen inte har undersökts eller kunnat undersökas tillräckligt noga.

Ett exempel är att den sjukpenninggrundande inkomsten, SGI, användes för kontroll av inkomststoppgift för bostadsbidrag trots att uppgifterna inte var jämförbara. Åtal väcktes i flera fall och visade sig sedan sakna grund (IRI-rapport 1984-2 refererad i RRV 1990).

Ett annat exempel från RRV:s studier av området, är att uppgifter i invandrarverkets SU-system inhämtades av polisen under icke kontorstid och användes i annat syfte än det avsedda. Det kunde ha lett till frihetsberövande på felaktig grund för enskilda personer (RRV 1985).

Ytterligare ett exempel är att data missuppfattas, så att fel information uppstår. Här kan det finnas många olika orsaker. Ett klassiskt IA-exempel är att begreppsbildning och namngivning inte är klarlagda. En annan möjlighet är att användaren bara har ytliga kunskaper om det område som data handlar om och att systemets presentation av data förutsätter en användare med expertkunskaper. Man kan t ex tänka sig att ordet "inkomst" i SGI-begreppet kan missleda den som inte känner till hela innebörden i SGI. Övriga avvikelser från god praxis i utformning av användargränssnitt kan också vara en orsak till feltolkningar.

Problemet kan även ha en omvänd effekt, nämligen att data med en för syftet god kvalitet inte används. Användaren/aktören kommer alltid att rätta sitt agerande efter *sin uppfattning* om kvaliteten på de data som finns tillgängliga! Användaren kan av olika skäl sakna förtroende för källan till data. Om kvaliteten inte är klargjord, t ex genom en deklaration, kan användaren bli tveksam om hur den information som finns tillgänglig ska användas. Användaren kanske då i stället fattar sitt beslut på ofullständig information eller lägger ned arbete på att själv söka fram andra data.

7.2 Deklarationens grundmaterial

Grundmaterialet till deklARATIONEN består i huvudsak av två delar. Den ena delen är en dokumentation av avsedd kvalitet som också inkluderar syftet med datainsamlingen. Den andra delen är uppgifter om faktisk kvalitet som tas fram genom kvalitetsmätningar, studier av registeregenskaper osv. Ett sådant grundmaterial bör innehålla allt som kan vara relevant på något sätt men kommer troligen inte att användas som det är. Urval bearbetas till olika presentationsformer för olika mottagare och syften.

Användningsområden för deklARATIONEN – mottagare

Nedan räknas några olika användningsområden upp tillsammans med krav på deklARATIONEN i respektive fall:

DeklARATIONEN hör till ett avtal

DeklARATIONEN ska ingå som en bilaga till ett avtal eller i form av en överenskommelse om att överföra data mellan två eller flera parter. Formkraven styrs av att deklARATIONEN ska vara en del av ett avtal. Största möjliga klarhet fordras, vilket kräver t ex entydighet i de begrepp som används i informationsmodellen. En deklARATION med detta syfte ska vara skriftlig och dokumentet ska ingå som bilaga i avtalet. I bilagan till avtalet ingår den specificerade kvalitet som informationsleverantören förbinder sig att upprätthålla. Brister i förhållande till den utlovade kvaliteten kan enligt avtalet t ex leda till avdrag på ersättningen.

Ett exempel på vad som kan anges i ett avtal kan vara en viss maximal ledtid från händelsetiden till dess att uppgiften finns i databasen eller sänds till mottagaren.

Den information och de egenskaper som inte tas upp i avtalet ska betraktas som ospecificerade uppgifter.

Informationsleverantören kan och bör dessutom tillhandahålla upplysningar om faktisk kvalitet som man genom egna mätningar eller erfarenhet har eller får tillgång till. Mottagaren kan då bli bedöma vilka egna kontroller av data som måste göras.

DeklARATIONEN vänder sig direkt till användaren

Det andra användningsområdet eller syftet är att ge användare som tar emot information i t ex en beslutssituation en så god bild som möjligt av informationens användbarhet i ett visst syfte. DeklARATIONEN vänder sig direkt till informationsanvändare inom den egna organisationen men kan också vända sig till andra användare. Formen styrs av syftet att användaren ska nås av budskapet – och förstå det. Formfrågorna kan t ex handla om vilken pedagogik som används, deklARATIONENS uppställning, språkbruk osv.

Med syftet för ögonen, kan man överväga olika sätt att distribuera deklara-tionsuppgifterna. Det är givetvis en fördel om så mycket som möjligt kan delges användaren via systemet självt, genom speciella datafält, ledtexter, hjälptexter osv. Möjligheten att användaren verkligen tar del av deklara-tionsinformationen ökar då. En deklaration i t ex skriftlig form under en flik i användarhandboken eller i ett produktblad som lämnas till den person som har ansvar för att anskaffa informationen till organisationen, skulle kanske inte uppmärksammas på samma sätt.

Deklarationen vänder sig till blivande användare eller informationsanskaffare

Detta är en deklaration som ges till en potentiell användare, dvs en använ-dare som ännu inte bestämt sig för att använda en viss informationskälla. Det kan röra sig om beskrivande produktblad för en viss databas eller en informationstjänst.

Deklarationen vänder sig till systemutvecklare och till förvaltningsansvariga (dataadministratörer etc)

Dessa personer kan dels vara med och bedöma om informationen är lämp- lig att användas i den egna organisationen, dels studera en deklaration för att avgöra hur informationen ska behandlas i de egna systemen.

Deklaration eller beskrivning är en grund för förändring och förbättring

Ett helt annat slag av användningsområde är rapportering av en mätning som gjorts i syfte att förbättra något. Mätningen kan t ex innebära att man går igenom ett register i syfte att finna systematiska fel vars orsak man kan åtgärda. En sådan mätning behöver kanske inte ge helt tillförlitliga mått på hur stora felen är, eftersom syftet inte är att redovisa felen utan att före- bygga dem. Att redovisa felen har dessutom ett begränsat värde, eftersom man ändå snart kommer att minska dem eller få bort dem helt. Sättet att mäta inriktas i detta fall på att hitta felkällor och felorsaker. Redovisningen får därför troligen en form som är mindre intressant för en användare. Användaren är säkert mer intresserad av en redovisning av de fel som på- verkar syftet med användningen, än en som detaljerat redovisar vad felen beror på.

Deklarationens innehåll och form

Vad ska tas med i deklARATIONEN?

Vilka uppgifter som är lämpliga att ta med varierar från fall till fall beroende på syftet med informationssystemet, syftet med deklARATIONEN, vilken slags data det rör sig om, vad data är avsedda att beskriva, hur mätning eller insamling har gjorts osv.

I första hand bör kraven på deklARATIONEN grundas på en riskbedömning som görs utifrån primäranvändarens situation.

Vilken ambitionsnivå är lämplig?

När man gör en deklARATION måste också en ambitionsnivå fastställas. Den följande framställningen omfattar ett rätt så stort antal punkter. Det kan finnas goda skäl att inte ta med alla uppgifter, utan göra ett urval. Urvalet görs med tanke på mottagarens behov, förkunskaper, referensramar och de risker man vill minska genom att ge mottagaren upplysningar om faktisk kvalitet. I avtalssammanhang får parterna komma överens om vilka faktorer som ska vara med och hur de ska beskrivas.

När deklARATIONEN riktar sig direkt till användare, kan urvalet baseras på en riskanalys som utgår från risken att användarna/aktörerna fattar felaktiga beslut med sakligt viktiga eller ekonomiska konsekvenser därför att de inte känner till vilken kvalitet informationen har.

Deklarationens uttryckssätt

- DeklARATIONER som vänder sig direkt till slutanvändare.

En vanlig användare ska inte behöva vara väl insatt i den terminologi som används av fackfolk för att beskriva informationskvalitet. I en deklARATION, som vänder sig direkt till användare, bör uppgifterna därför i första hand uttryckas i klartext.

Hur ska man då uttrycka sig? I stället för att säga att aktualiteten är 14 dagar, kan man t ex säga att det i genomsnitt tar 14 dagar från det att ett företag ändrar telefonnummer till dess att ändringen finns införd i registret, maximalt tar det sex månader och i sju fall av tusen kommer ändringen inte alls in. Den senare formuleringen har fördelen att den är mer entydig, och även kan tolkas rätt om den tas ur sitt sammanhang.

Antag att man också vill säga något om ändringstakten för de olika uppgifterna. Man kan t ex skriva följande:

"Under ett år ändras i genomsnitt uppgiften om telefonnummer i registret för 11% av de företag som finns i registret."

eller

"Uppgiften 'ägare' ändras under ett år i genomsnitt för 3% av de företag som finns i registret."

DeklARATIONEN kan också mycket väl anges i kort form med respektive faktor. Det kan till och med vara en fördel. Läsaren lär sig då terminologin och får på så sätt lättare att läsa andra deklARATIONER som inte har en användaranpassad terminologi.

- Deklarationer som vänder sig till specialister som ska ta ställning till vilken information som ska anskaffas eller som ska ingå i ett avtal.

De olika uppgifterna placeras under rubriker, som överensstämmer med de olika egenskaperna (tid, aktualitet, noggrannhet osv). Själva uppgiften kan mycket väl uttryckas i klartext, under förutsättning att det ger en entydig beskrivning.

Det är ingen nackdel om uppgifterna uttrycks både i klartext och med hjälp av facktermerna:

”Noggrannheten är fem procent, dvs de angivna värdena ligger i genomsnitt inom plus, minus fem procent från ett ’sant’ värde.”

Vad ska beskrivas, fakta eller bedömningar?

I en deklARATION eller en beskrivning av data, kan både en värderande och en faktabeskrivande ansats ingå. Ansatserna bör dock klart skiljas åt eftersom informationsproducenten och mottagaren kan ha olika uppfattning om hur egenskaper hos data påverkar lämpligheten för ett visst syfte.

En värderande ansats innebär att informationsgivaren uttalar sig om datas värde för ett visst syfte och för en viss tänkt användare eller användargrupp. Deklarationen kan t ex innehålla ett uttalande om en uppskattad tidsgräns för användbarheten hos data. Tidsgränsen relateras då till krav på noggrannhet och på kunskap om hur snabbt variabler ändrar sig.

Den faktabeskrivande ansatsen innebär att så mycket grundfakta som möjligt lämnas om informationen. Dessa grundfakta bör då vara uttryckta så att användaren kan relatera dem till sina kvalitetsvärderingar och göra en egen bedömning. Med denna inriktning bör man även lämna uppgifter om hur relevanta förhållanden och därmed variabler ändrar sig m m. Det ger användaren möjlighet att göra en egen bedömning av tidsgränsen för användbarhet. Man kan exempelvis lämna uppgifter om hur många nya objekt som registreras i ett register en genomsnittlig månad.

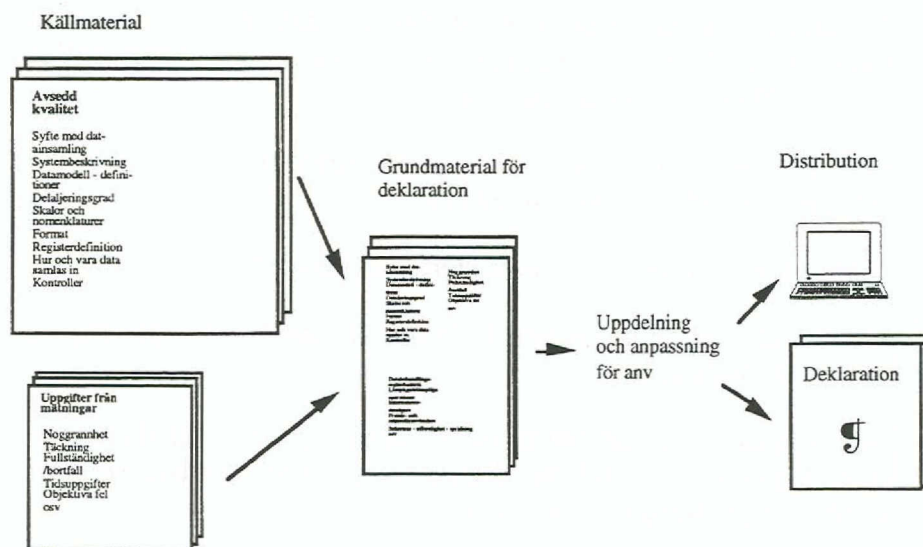
Som användare vet vi aldrig säkert vilka faktorer producenten har vägt in i sin bedömning av informationens användbarhet för ett visst syfte. Visste vi det så kunde vi kanske lika gärna själva göra bedömningen utifrån grundfakta. Vi måste själva ta ställning till om vi har alla grundfakta vi behöver. Vår bedömning av producentens bedömning får grunda sig på de fakta vi själva känner till, vår kännedom om producenten och vår uppfattning om producenternas trovärdighet (som kan vara både för hög och för låg), beroende på tidigare erfarenheter.

Ett exempel: Ett grundfaktum kan vara vilken tidpunkt en viss uppgift gäller. Det kan kallas händelsetid:

”Anders Andersson lämnade in ändringsanmälan om ny adress den 23/4 1992.”

Det användningsrelaterade begreppet är i detta fall aktualitet. Användaren gör en bedömning av om uppgiften är tillräckligt aktuell för att kunna användas i det aktuella syftet. I denna bedömning kan ytterligare fakta vara användbara, t ex att användaren vet att Anders Andersson flyttar regelbundet vart tionde år. Detta kan sägas vara ett specifikt fall.

Om en användare i stället ska bedöma ett typfall, blir situationen något annorlunda. Anta att användaren ständigt tar emot stora mängder adressändringar. I en sådan situation kan grundfakta röra prestanda hos informationsprocessen, t ex hur lång tid det i genomsnitt tar från det att en händelse inträffar till dess att informationen är tillgänglig för användaren. Användaren får göra en bedömning av den stora massan ärenden, och om t ex någon extra kontroll behövs i specifika fall.



Figur 8 – Deklarationens sammanställning och spridning.

Deklarationens förvaltning

Förvaltning av deklarationen ingår som en del av det förvaltningsarbete som DA- eller IA-funktionen utför. Förändringar och förbättringar leder till att innehållet i deklarationen måste ändras. Brister i datakvalitet när det gäller data om data kan vara lika förödande som brister i data.

Kopplingen mellan verksamhetsmodell, datamodell och deklaration måste vara given och bör helst vara maskinellt kontrollerad så att inte inkonsistens uppstår mellan dessa tre eller mellan olika verksamhetsområden.

Deklarationen kan i vissa fall behöva registreras så att den inte kan ändras okontrollerat, t ex om den är kopplad till ett avtal. Ändringar får då bara göras enligt de former som stadgas i avtalet.

Epilog

”Jaha Nu är serverprogrammet för Spegelbasen uppdaterat. Då ä bar å åk! Lycka till.”

Svenne lämnar rummet.

Du slår dig ner framför PC:n och startar programmet. Data hämtas in från olika håll allt eftersom du frågar.

Aha! Där blinkar en röd signal på statistiken över installerade maskiner. Bäst att se efter vad det betyder.

Ej kontrollerad siffra. Kontrollerade data kommer om ca 14 dagar. Installerad mängd brukar i okontrollerade data överstiga korrekt värde med 7–9 %. Observera att följande ej ingår enligt registerdefinitionen: *OEM-maskiner*

Blå signal på siffrorna över försäljning tidigare år – historiskt säkra och kontrollerade data med andra ord.

Orange signal på siffran för inversion. Vad betyder det då?

Välkontrollerad online-uppift, har varierat med 5–20 % per dag de senaste 20 dagarna.

Och grön signal på uppgifterna om köparna, Oj! rena plantskolan, sju olika källor till den bilden. Bäst att gräva lite mer och se efter varifrån uppgifterna kommer. Aha, hm...

– Nåja, det duger.

Nu är det bara att kolla vilka andra som har tillgång till dom här siffrorna. Mainiaderindex sprids till alla i branschen, hm. Men siffrorna över Krimfabrikernas produktion är det bara vi och Anderssons som har just nu i Sverige! Nu förstår jag varför dom...

Din tillförsikt ökar. Nu är det dags att agera.

Några ordförklaringar

DA - DataAdministration. Arbete med att ordna och beskriva data inom en verksamhet.

Posten definierar DA så här:

”Dataadministration innebär att *medvetet* åstadkomma entydigt beskrivna, ensade och stabila begrepp som är tillgängliga på lämpligt sätt och som skapar förståelse för de data som används för att avbilda Postens verksamhet.

Detta sker genom att använda i Posten vedertagna metoder, tekniker och verktyg för att i samverkan kartlägga, strukturera, beskriva och presentera data och datautbyte för/med systemutvecklare, systemförvaltare och användare.”

Se vidare om dataadministratörens arbete i TRIAD V4, sidan 51.

Datakatalog (data dictionary) – Förteckning över datatermer med betydelse angiven. Bygger på en viss datamodell. Kan även innehålla attribut, samband, operationer m m. Kan implementeras i ett DKS - DataKatalogSystem.

IA, informationsadministration – IA presenteras på följande sätt i TRIAD rapport V4:

”Informationsadministration (IA) – Aktiviteter som syftar till att:

- underlätta informationsbildning. IA beskriver relevanta begrepp och data i en verksamhet för att aktörerna ska kunna tolka data på ett entydigt sätt
- främja effektiv dataåtkomst. IA ser till att data finns och kan nås på ett för aktören relevant sätt.”

IRDS (Information Resource Dictionary System) Resurskatalog kan lite informellt beskrivas som ett mer avancerat och påbyggt datakatalogsystem. Se dock övriga TRIAD-rapporter i K-serien för bättre beskrivningar! IRDS beskrivs i TRIAD rapport K1 så här:

”Ett IRDS adresserar främst tre områden:

- den ger en flexibel datamodell
- den svarar för generella tjänster som versionshantering och säkerhet
- den ger informationsgränssnitt i form av skärmbilder för slutanvändare, brygger till andra resurskataloger samt åtkomst av data för CASE-verktyg och tillämpningar.”

Referenser

Andersson Bengt E W, Nilsson Peter, *How to manage complexity in inter-organizational information systems (IOIS) – some preliminary conclusions*. Bidrag till IFIP WG8.5 working conference on Decision Support in Public Administration. Delft, Holland 13 - 14 May 1993

Bergman Bo, Klevsjö Bengt, *Kvalitet från behov till användning* Studentlitteratur. 1991

Brammer Mogens, Danish Standards Association: *Data Quality, Meaning of the concept and recommended actions*, To CEN/TC3311 meeting 13-14 sept 1993.

Crosby Philip R, *Quality is free The art of making quality certain*. Mentor executive library, 1980 .

de Bono Edward, *Tänketräning*. Svenska Dagbladets Förlags AB, 1984

Eneroth Bo, *Hur mäter man "vackert"?* Grundbok i kvalitativ metod Natur och Kultur, 1992.

Institutet för kvalitetsutveckling, *Kriterier och anvisningar för utmärkelsen svensk kvalitet 1993*

Ledell Göran (red), *Att mäta informationens användbarhet*. Dataföreningen i Sverige, 1993, Studentlitteratur, Lund.

Nilsson Peter, *Information Quality. An analysis of quality aspects of information used in problem solving*. Lic uppsats, Institutionen för Data- och Systemvetenskap Stockholms universitet/Kungliga Tekniska Högskolan Stockholm, 1990.

Nørretranders Tor, *Märk världen* Bonnier Alba, 1993.

Peters, *Skapande kaos företagets revolutionshandbok* Svenska Dagbladets Förlags AB, 1988

Posten, *Dataadministration och EDI i Posten - en översikt*. Posten, Informatiksekretariatet vers 10UX, 1993.

Price Frank, *Kvalitet i alla led, Demingmetoden* SVD förlag, 1992

RRV, *ADB-säkerhet – Granskning av de civila myndigheternas ADB säkerhetsarbete*, 1985.

RRV, *Rätt data? Studier av informationskvalitet i statliga ADB-system*. Revisionsrapport, 1990

RRV, *ADB i samverkan. En granskning av ADB-utvecklingsprojekt med flera myndigheter som deltagare*, 1991

RRV, *Fel data kostar! Exemplet Postens kostnader för fel i adressregister*, 1992

SFK-DATA, **Sveriges Verkstadsindustrier**, VI, *TickIT-Handboken Handledning vid konstruktion och certifiering av kvalitetssystem för programvara med användning av SS-ISO 9001*. Utgåva 1, januari 1993.

ITS, *Terminologi för informationssäkerhet*. Rapport ITS 6 Utgåva 1 Mars 1994.

Anm. Detta är efterföljaren till den tidigare **SIS** teknisk rapport 322, *Informationssäkerhet och dataskydd – en begreppsapparat* 1989.

Socialstyrelsen, folkhälsenheten, epidemiologiskt centrum, och Uppsala Universitet, Socialmedicinska institutionen, *Kvaliteten i Socialstyrelsens slutenvårdsregister för 1986, kort sammanfattning av vissa resultat*. PM, 1992.

Statskonsult Norge, *Nasjonal infrastruktur for edb, Prosjekt 5 Standardisering av formidlingskanaler, organisering og styring av formidlingen av offentlig elektronisk informasjon*, 1991.

Svensson Roy, Östling Anna Karin, *Databaser Ett skadeståndsrättsligt perspektiv*. Institutet för Rättsinformatik IRI-Rapport 1992:5.

TRIAD-projektet, Telia, Posten, Statskontoret, Ericsson och SISU, Marianne Janning, Clary Sundblad. *Modellkvalitet*, TRIAD-rapport nr 7/93.

TRIAD-projektet, Telia, Posten, Statskontoret, Ericsson och SISU, Harriet Borgman *Namnsättning i modelleringssammanhang*. TRIAD Modelleringshandboken N 10:11

TRIAD-projektet, Telia, Posten, Statskontoret, Ericsson och SISU. Christina Johannesson, Christer Dahlgren, Peter Söderström. *Den gemensamma informationsmarknaden. En referensram för handlingsfrihet och konkurrenskraft*, 1993.

Teldok, *Rapport 85 Vård och råd på tråd*. Reportage av Mats Utbult om distansdiagnostik och telemedicin som stöd för samverkan och lärande i sjukvården – från Nordkalotten till Skåne-Blekinge, 1994

Televerket, *Datakvalitet i själva verket (Rätt på direkten!) - en slutrapport*, 1991.

Wurman, R S *Information Anxiety*, 1989 Doubleday, New York.

TIDIGARE UTGIVNA PUBLIKATIONER AV TRIADGRUPPEN

Verksamhetskrav på informationsadministration

- V 1: IA och verksamhetens krav – erfarenheter från offentlig förvaltning
- V 2: Fallstudie av IA-projektet vid Televerket
- V 3: IA-erfarenheter från företag och myndigheter
- V 4: Den gemensamma informationsmarknaden – en referensram för handlingsfrihet och konkurrenskraft
- V 5: ...fråga är guld. Lokal affärsstyrning utifrån den egna verksamhetens data
- V 6: Det kundeffektiva företaget

Modellering

- N 1: Modelleringsansatser för begrepps- och datamodellering – Beskrivning och försök till jämförelse
- N 2: Generering av konceptuella modeller från policydokument
- N 3: Espritprojektet Tempora
- N 4: Prövning av regelbaserad metodik inom Posten
- N 5: En kokbok i remodellering – utkast
- N 6: Datorstöd för modellintegration
- N 7: Modellbaserad kunskapsinsamling
- N 8: Modellkvalitet
- N 9: Samband mellan dokument och modeller
- N 10: Modelleringshandboken
 - 1 – Översikt
 - 2 – Modelleringsledarens bashandledning
 - 3 – Modellering i grupp
 - 4 – Kommunikation
 - 5 – Arbetsgångar
 - 6 – Modelleringsväskan
 - 7 – Objektorienterad verksamhetsanalys
 - 8 – Basmodeller
 - 9 – Regelmodellering i praktiken
 - 10 – Business Process Reengineering
 - 11 – Namnsättning
 - 12 – Tolkning av grafiska modeller
- N 11: Ett+Ett=Ett – Två praktikers erfarenheter av modellintegrering

Kunskapsförmedling

- H 1: Handedarutbildning för modelleringsledare, avancerad
- H 2: Slutrapport HUMLA prototyp
- H 3: Utbildning i Informationsadministration
- H 4: Spridning av Hybris – en fallstudie vid Telia

Uttagssystem

- U 1: Hybris i Unix-miljö
- U 2: DEBRIS
- U 3: Hybris DOS/PimWin på Posten
- U 4: Program för sökning i databaser – en marknadsöversikt
- U 5: Att nå och förstå data – möjligheter och begränsningar

Katalogprinciper

- K 1: IRDS
- K 2: IRDS Modeller och modellnivåer
- K 3: Koppling begreppsmodell – relationsmodell
- K 4: IBM:s Repository Manager – en introduktion
- K 5: IBM:s Repository Manager: Datamodelleringsbegreppen
- K 6: IBM:s Repository Manager: Begreppsmodellering i Information Model
- K 7: IBM Repository Manager: Attribut- och värdemodellering i Enterprise Submodel
- K 8: Navigering i Repository
- K 9: TRIAD Newsletter – IRDS inom ISO. Dagsläget
- K 10: TRIAD Newsletter – ISO/IRDS. Händelseutvecklingen 91/92
- K 11: Samverkan mellan resurskataloger – visioner eller behov
- K 12: AD/Cycle 1 Information Model – Processer och informationsflöden mellan processer
- K 13: AD/Cycle 1 Information Model – Info Flows inom Processmodellen
- K 14: AD/Cycle 1 Information Model – Relationsdatabasmodellering
- K 15: AD/Cycle 1 Information Model – Härlednings-specifikationer i begreppsmodellen
- K 16: IA-prototyp
- K 17: Repository AD/Cycle – International Users Group
- K 18: RAD-konferensen i Chicago, 1992
- K 19: Vad händer inom ANSI-IRDS?
- K 20: Information Warehouse – vad är det?
- K 21: CDIF – en översikt
- K 22: PCTE – en översikt
- K 23: XLII – en öppen och flexibel utvecklingsmiljö
- K 24: Hybris IA/DA – En IA-prototyp vid Telia
- K 25: Introduktion till GDMO-standarderna
- K 26: OpenODB
- K 27: ANSI/X3H7 "Object Information Management"
- K 28: Object Management Group
- K 29: Översättning av modelldata
- K 30: Objektorienterade ansatser inom ANSI/IRDS
- K 31: Repository, State-of-the-Art

KORT OM TRIAD

Triad är namnet på ett treårigt samarbetsprojekt kring informationsadministration och dataadministration, IA/DA, som Telia, Posten, Ericsson, Statskontoret och SISU bedriver. Syftet är att utveckla parternas synsätt, metoder och hjälpmedel inom detta område.

Arbetet inom Triad är uppdelat i delprojekt som är sammanförda i tre block.

Beställarblocket vänder sig dels till dem som är verksamhetsansvariga och måste ta ställning till IA-/DA-satsningar, dels till dem som har ansvaret

för IA/DA inom en organisation. Delprojekten inom detta block arbetar med att formulera verksamhetens krav på IA/DA samt studerar och beskriver roller, organisation och arbetsformer för IA-/DA-arbete.

Utförarblocket vänder sig till dem som arbetar med IA/DA.

Delprojekten arbetar med modellering, data- och resurskataloger samt uttagssystem.

Kunskapsförmedling är det block som ser till att resultaten kommer Triad-parterna till godo. Detta sker bland annat genom kurser, seminarier samt genom att rapporter som denna ges ut.
