

$$1 + 1 = 3?$$

# NYTTOR OCH UTMANINGAR MED ÖPPNA LÄNKADE FÖRETAGSDATA

DECEMBER 2012

PÄR LANNERÖ, METAMATRIX AB,  
[PAR.LANNERO@METAMATRIX.SE](mailto:PAR.LANNERO@METAMATRIX.SE)

JAN HYLÉN, IVAR H KONSULT AB,  
[JAN@JANHYLEN.SE](mailto:JAN@JANHYLEN.SE)

JOHAN GROTH, GROTH & GROTH AB,  
[JOHAN.GROTH@GOGAB.SE](mailto:JOHAN.GROTH@GOGAB.SE)

# INNEHÅLLSFÖRTECKNING

Innehållsförteckning.....	2
Sammanfattning.....	4
Inledning.....	4
Bakgrund och centrala begrepp.....	6
Data.....	6
Öppna data.....	6
Öppna offentliga data.....	7
Öppna företagsdata.....	8
Länkade data.....	8
Länkade öppna data.....	9
Länkade öppna offentliga data.....	10
Masterdatahantering.....	11
Nytta.....	12
Nyttor.....	12
Allmänna nyttor.....	12
Demokrati.....	12
Rättssäkerhet.....	13
Aktualitet och snabbhet.....	13
Kvalitet.....	13
Alternativkostnader.....	14
Tillgång till information.....	14
Sök- och hittbarhet.....	15
Mer fullständig och allsidig information.....	16
Innovationer.....	17
Vetenskaplig utveckling.....	18
Teknisk utveckling.....	18
Automatisering.....	19
Myndighetsspecifika nyttor.....	19
Kategorisering av nyttor.....	21
Ekonomiska beräkningar av nyttan med öppna data.....	22
Kostnader och risker.....	23
Utmaningar.....	24
Nyttan ligger hos någon annan.....	24
Brist på ekonomisk analys.....	24
Brist på information.....	24
Otydligt vem som äger informationen.....	25
Brist på tydlighet hur man får återanvända informationen.....	25
Data erbjuds inte i öppna format.....	25
Förutsättningar för publicering.....	25
Förutsättningar för användning.....	26
Kostnader.....	26

Allmänna risker .....	27
Myndighetsspecifika risker .....	29
Några tänkbara scenarion .....	29
Jag behöver grunduppgifter om en viss person. ....	29
Beskrivning .....	30
Fördelar med semantisk teknik .....	31
Utmaningar.....	31
Jag vill ändra mina uppgifter. ....	31
Beskrivning .....	32
Fördelar med semantisk teknik. ....	33
Utmaningar.....	33
Offentlig upphandling.....	33
Beskrivning .....	34
Fördelar med semantisk teknik .....	35
Utmaningar.....	35
Slutkommentar.....	35
Bilaga 1: Vad menas? - Terminologisk ordlista med definitioner .....	37

## SAMMANFATTNING

Öppna data kan leda till ökad insyn, innovationer och besparingar. Länkade data underlättar återanvändning och automatisering. Masterdatahantering ökar kvalitet och underlättar organisatoriska förändringar.

Denna rapport förklarar begreppen och på vilket sätt nyttorna kan realiseras men tar också upp ett antal utmaningar.

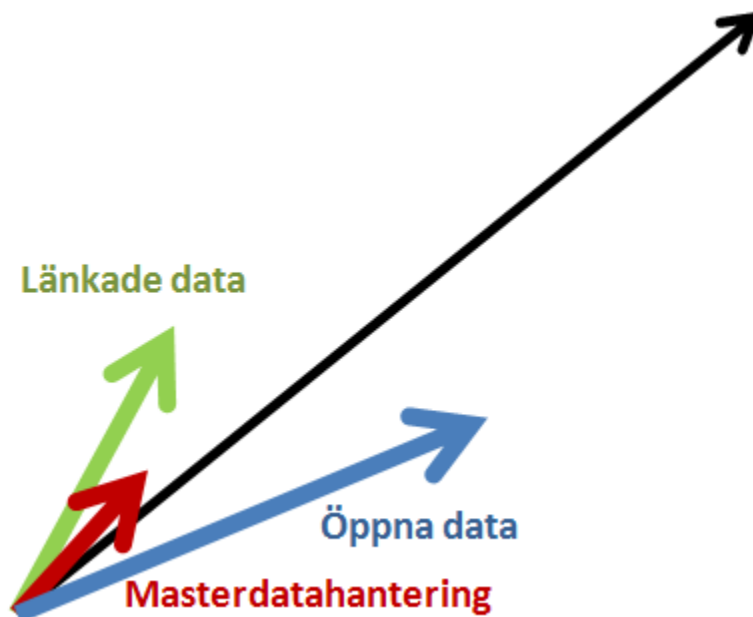
Den primära målgruppen är beslutsfattare som behöver orientera sig inför eventuella satsningar på öppna och länkade data. Vårt huvudfokus har varit offentlig sektor, men resonemangen är i stor utsträckning giltiga även i andra sammanhang.

Rapporten är en del av ett större utvecklingsprojekt som samordnas av Skatteverket och som omfattar ett antal myndigheter, organisationer och företag i Sverige och Norge. Rapporten blir härigenom en steg i ett bredare arbete som kommer att fortgå under kommande år.

## INLEDNING

Denna rapport är framtagen på uppdrag från Skatteverket Vinnovaprojektet "Öppna upp företagsdata" (Dn. 2012-01675). Uppdraget var att skriva rapport som på ett lättillgängligt sätt visar på nyttor och utmaningar med öppna länkade data, utan att läsaren ska behöva ha en djupare förkunskap. Författarna svarar ensamt för innehållet i rapporten.

Projektet är ett försök att konkret påvisa och realisera potentialen i tre trender, som diskuterats under flera år, men som nu börjar bli mogna att tillämpa i praktiken: Öppna data, länkade data och masterdatahantering.



*De tre trenderna har ungefär samma riktning. Kombineras de kan betydande nyttor realiseras.*

Projektet har valt att fokusera på grundläggande uppgifter om juridiska personer som vi valt att kalla för "företagsdata". Företagsdata är ett exempel på värdefulla data som ofta återanvänds, ofta har meningsfulla relationer till andra data, sällan är sekretessbelagda och som dessvärre ofta finns i olika, sinsemellan motstridiga versioner. Därför tror vi att de erbjuder "långt hängande frukter" för den som vill dra nytta av de tre trenderna.

Nyttorna är i stor utsträckning generaliserbara till andra typer av data som myndigheter hanterar varför denna rapport har ett värde även utanför det specifika området företagsdata.

Rent konkret handlar det om att de inblandade myndigheterna på ett systematiskt sätt gör företagsdata tillgängliga för återanvändning:

- på internet
- maskinläsbart
- med standardiserade identifierare
- väl dokumenterat
- med hög kvalitet, fullständighet, beständighet och säkerhet

Förutsättningarna för detta finns till delar inom svensk förvaltning med offentlighetsprincipen, en relativt höga IT-mognad etc. men utmaningar finns det likväl kvar, framför allt vad gäller informationsmodeller, begrepp och formatstandarder.

Vi inleder med en presentation av de tre trenderna, och förklarar några andra i sammanhanget viktiga begrepp. Därefter redovisar vi potentiella nyttor, kostnader och utmaningar. I en förhoppningsvis kommande rapport kommer även ett antal fallstudier att redovisas.

## BAKGRUND OCH CENTRALA BEGREPP

I detta kapitel ges en introduktion till centrala begrepp. Som en del i projektet Öppna upp företagsdata har terminologiska begreppsanalyser gjorts av de centrala begrepp som figurerar i denna rapport. I bilaga 1 finns en terminologisk ordlista med de termer och definitioner som projektgruppen står bakom.

### DATA

I Rikstermbanken<sup>1</sup> finns följande definition av data: ”Representation av fakta, idéer eller liknande i en form lämpad för överföring, tolkning eller bearbetning av människor eller av automatiska hjälpmedel”.

Vidare: ”Data blir information när någon har tolkat innebörden av data. Många gånger behöver man inte hålla isär begreppen data och information. Men exempelvis vid överföring mellan datorer eller lagring i datorminnen är det data, inte information, som hanteras.”

### ÖPPNA DATA

Öppna data är data som görs digitalt tillgängliga för att underlätta återanvändning av vem som helst - ofta för andra ändamål än de ursprungliga.

De flesta menar dessutom att öppna data måste vara gratis, och vissa menar att även andra trösklar - såsom begränsande val av filformat - måste slipas ner för att etiketten ”öppna data” ska få användas.

Det finns en handfull föreslagna definitioner, från mer eller mindre officiella källor. Här är några:

”Data anses vara öppna om det står var och en fritt att använda, återanvända och sprida med angivande av källa och/eller krav på att dela på samma villkor som största motprestation.”<sup>2</sup>

”Any sets of data which can be reused with no restrictions by any form of licensing or patents, data that are well structured and can be easily accessed and reused by institutions, scientists or the web community”<sup>3</sup>

”Komplett, primär, aktuell, tillgänglig, maskinläsbar, fri, i ett öppet format”<sup>4</sup>

---

<sup>1</sup> <http://www.rikstermbanken.se/> Källa: Terminologicentrum TNC: Basord i våra fackspråk (2012)

<sup>2</sup> <http://opendefinition.org/> i översättning av Peter Krantz

<sup>3</sup> <http://www.slideshare.net/peterkz/open-government-data-on-the-web-a-semantic-approach>

<sup>4</sup> <http://www.opengovdata.org/>

Oavsett val av definition är det uppenbart att många tror att öppna data kan realisera stora värden, både för samhället i stort och för enskilda. Listan med portaler, konferenser, arbetsgrupper och programmeringsevenemang som fokuserar på öppna data kan göras lång.

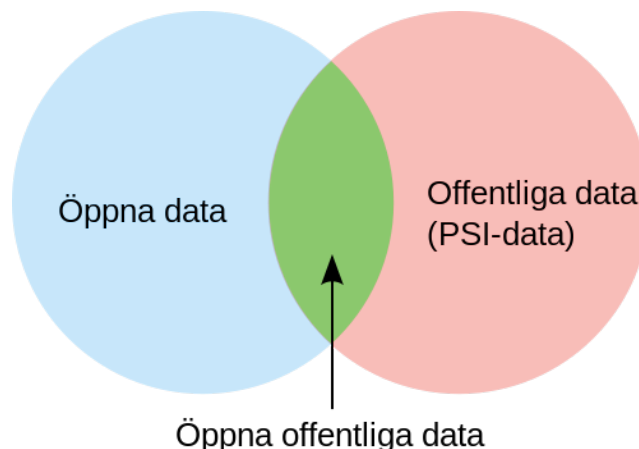
Språkrådet, TNC och övriga i Datatermgruppen är överens om det som nu står på Datatermgruppens webbplats<sup>5</sup>:

”Data är ett pluralt ord. Det kan ha den mer konkreta betydelsen ’uppgifter’: känsliga data, geografiska data. I it-sammanhang kan det ha en mer kollektiv, abstrakt innebörd (jämför information, kosmetika) och behandlas då ibland som ett singulart massord: tillgänglig data, all data, öppen data (med n-genus). Vi rekommenderar dock att man fortsätter att behandla data pluralt: dessa data, öppna data. När man vill framhäva den kollektiva innebörden bör man hellre precisera sig, t.ex. genom att i stället tala om datamängd.”

## ÖPPNA OFFENTLIGA DATA

Med öppna offentliga data (ibland *öppna myndighetsdata*) avses öppna data från myndigheter eller andra offentliga institutioner. EU har under en längre tid varit en stark förespråkare för att tillgängliggöra offentliga handlingar (eng. Public sector information - PSI) som öppna data. Redan 2003 kom PSI-direktivet<sup>6</sup> som med vissa begränsningar förpliktigar europeiska myndigheter att göra sina handlingar tillgängliga för återanvändning.

I USA har Obama-administrationen<sup>7</sup> vidtagit snarlika åtgärder, och runtom i världen finns många exempel<sup>8</sup> på liknande satsningar. Exempelvis har den svenska regeringen nyligen beslutat starta portalen [opnadata.se](http://opnadata.se).



<sup>5</sup> <http://www.datatermgruppen.se/>

<sup>6</sup> Europaparlamentets och rådets direktiv 2003/98/EG om vidareutnyttjande av information från den offentliga sektorn

<sup>7</sup> [http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda\\_2010/m10-06.pdf](http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf)

<sup>8</sup> <http://www.data.gov/opendatasites> är en bra startpunkt





*Fler och fler databaser länkas samman på webben, bilden visar några av dem  
(CC-BY-SA Richard Cyganiac)*

Den traditionella webben består inte av länkade data utan av länkade dokument och den som är bekant med databaser inser att det finns en del fördelar med länkade data i jämförelse med länkade dokument: det går mycket lättare för program att automatiskt sortera, analysera, återanvända och bearbeta data än att göra detsamma med dokument.

Det är t.ex. svårt att be Googles sökmotor att ta reda på hur många företag som finns i Sundsvall och hur många anställda de har i genomsnitt, men fullt möjligt att be ett program analysera länkade företagsdata och redovisa svaret på de frågorna.

Alla har inte läs- och skrivrättigheter till alla delar av den framväxande globala databasen av länkade data, men en stor del är helt öppen, och vem som helst kan bidra till att utöka den. Det som krävs är att vi skapar globalt unika identifierare (så kallade URI:er) för nya resurser och använder befintliga identifierare när vi refererar till resurser som redan är kända i den globala databasen.

Genom att man använder samma identifierare när man refererar till samma sak blir det betydligt enklare att *sammanställa* data från flera källor än när man arbetar med icke-länkad data.

Om vi till exempel vill publicera information om Sundsvall som länkade data bör vi alltså söka upp och använda en URI som identifierar kommunen. Till exempel <http://dbpedia.org/resource/Sundsvall>. Sedan underlättar det förstås om vi använder standardiserade modelleringsformat, kommunikationsprotokoll m.m. för att publicera våra data. The World Wide Web Consortium (W3C) har tagit fram en hel svit av sådana tekniska standarder<sup>9</sup>, inklusive frågespråket SPARQL, datarepresentationsspråket RDF och begreppsmodellsspråket OWL.

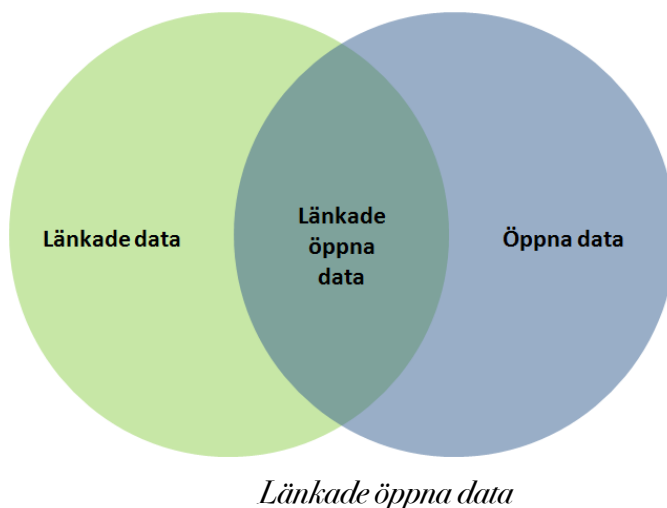
Tidigare talade man om *den semantiska webben*, men det uttrycket är nästan övergivet, eftersom det svåra ordet ”semantik” har en tendens att försvåra kommunikation utanför en begränsad skara av akademiker.

## LÄNKADE ÖPPNA DATA

Snittmängden mellan länkade och öppna data är särskilt intressant ur ett återanvändningsperspektiv eftersom öppenheten gör data tillgängliga och länkningen gör dem enkla att kombinera.

---

<sup>9</sup> <http://www.w3.org/standards/semanticweb/data>



Webbens skapare Sir Tim Berners-Lee har föreslagit en indelning i fem nivåer för öppna data<sup>10</sup>, där återanvändbarhet och nytta ökar för varje steg. På de högre nivåerna kan vi prata om *Länkade öppna data* (på engelska Linked Open Data - LOD), vilket för oss närmare visionen om en semantisk webb:

Öppna data i godtyckligt format...

... i strukturerade format

... i icke-proprietära format

... som har URI:er som nycklar

... som är sammanlänkade med andra öppna data på nätet

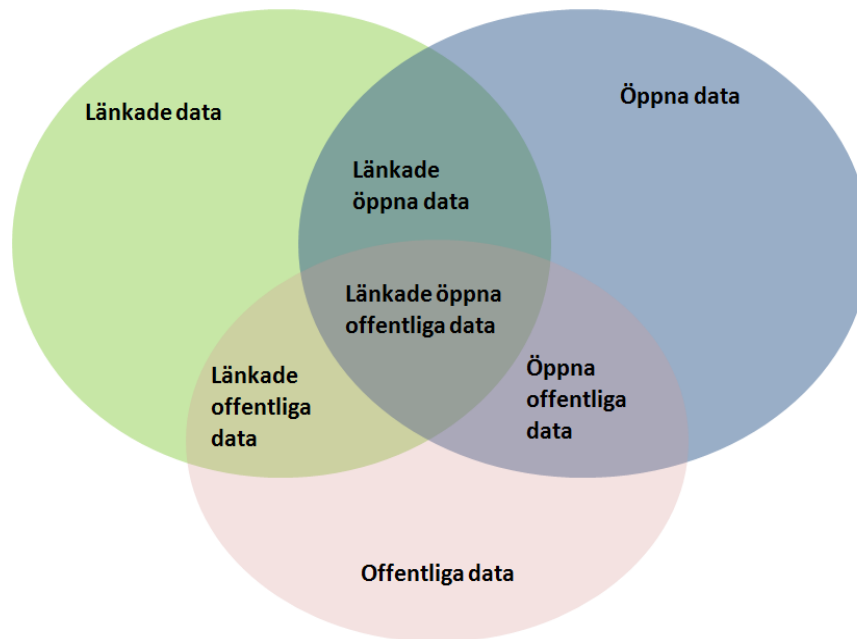
Man pratar om att femstjärnig data inte är *data på webben* utan *data i webben*, vilket hänvisar till den globala databasen.

## LÄNKADE ÖPPNA OFFENTLIGA DATA

Representanter för offentlig sektor utgör, kanske inte helt förvånande, en betydande del av de aktörer som engagerat sig kring länkade öppna data. Det finns till och med en särskild arbetsgrupp inom webbstandardiseringsorganisationen W3C som handlägger just länkade öppna offentliga data. Den kallas Government Linked Data Working Group<sup>11</sup> och arbetar mycket aktivt med att sprida best practices m.m.

<sup>10</sup> <http://lab.linkeddata.deri.ie/2010/star-scheme-by-example/>

<sup>11</sup> <http://www.w3.org/2011/gld/>



*Här är det område vi rör oss inom.  
Observera att företagsdata kan förekomma var som helst i bilden.*

## MASTERDATAHANTERING

Nästan varje organisation har en handfull databaser där uppgifter om kunder, medlemmar, anställda, produkter m.m. lagras. Ofta överlappar databaserna varandra och det är inte alltid säkert att alla databaser har aktuella uppgifter. Risken finns att den ena databasen uppdateras men inte den andra, vilket kan leda till förvirring, felaktiga utskick, duplicerade poster m.m. Och det blir inte lättare när organisationer slås samman eller delar på sig.

En lösning kan vara att bryta ut sådana data som förekommer i många olika system till en central masterdatabas. Så kan t.ex. adressuppgifter delegeras till en adressdatabas istället för att modelleras in i varje verksamhetssystem. Detta är en av flera strategier för *masterdatahantering*<sup>12</sup> (eng. master data management). Vill man undvika tekniskt klingande termer kan man istället tala om *en enda säker källa*.

Masterdatahantering diskuteras ofta från ett organisationsperspektiv. ”När [masterdata] saknas, är felaktig eller svår att hitta lider organisationerna av låg effektivitet, förseningar, fel och fattar sämre affärsbeslut”, skriver t.ex. Nathaniel Rowe, undersökningsanalytiker för företagsdatahantering hos Aberdeen Group<sup>13</sup>. Samma grundprinciper gäller, menar vi, även för ”organisationen” den offentliga sektorn, och för samhället i stort.

<sup>12</sup> <http://msdn.microsoft.com/en-us/library/bb190163.aspx>

<sup>13</sup> <http://www.stibosystems.com/Default.aspx?ID=5868>

## NYTTA

En nytta kan definieras som *resultatet av en förändring som upplevs som positivt av dess intressenter*, skriver E-delegationen i sin ”Vägledning till nyttorealiserings”<sup>14</sup>.

## NYTTOR

I detta avsnitt beskriver vi först olika typer av nyttor som vi bedömer att öppna länkade data har potential till. Resonemangen är inte specifika för företagsdata utan mer övergripande. Vi inleder med allmänna nyttor och fortsätter med myndighetsspecifika. Därefter kommer en mer detaljerad kategorisering av de olika nyttorna utifrån vem de berör, vilken trend de går att härleda till och huruvida de går att kvantitativt värdera.

Just frågan om vem som gynnas är av särskilt intresse, eftersom det med öppna data ofta är så att nyttan inte alltid finns hos den som öppnar upp sina data. Det är inte ovanligt att en stor potentiell nytta har samhällsekonomisk karaktär, men förutsätter att en enskild myndighet eller organisation bär en kostnad. I sådana lägen är det förstås viktigt att beslut och resursfördelningar görs från centralt håll.

Kapitlet avslutas med ett avsnitt om de ekonomiska aspekterna av öppna länkade data.

## ALLMÄNNA NYTTOR

### DEMOKRATI

Öppna myndighetsdata kan lyfta begreppet ”offentliga handlingar” till en ny nivå. En handling kan vara offentlig till ingen nytta, om ingen medborgare vet om att den finns och begär ut den. Men om öppna myndighetsdata görs maskinläsbara kan de *systematiskt* hämtas ut och *automatiskt* granskas.

Detta ger ökade möjligheter för journalister och andra att hitta förbättringsmöjligheter och missbruk, för lärare att skapa visualiseringar som ger förståelse osv. Medborgare som har tillgång till fullständiga rådata har på ett nytt sätt möjlighet att vara demokratiskt delaktiga. Begrepp som transparens och data journalism hör också hemma här.

I vissa länder<sup>15</sup> som saknar den tradition av öppenhet som offentlighetsprincipen innebär införs öppna data på grund av de samhällsekonomiska fördelarna. Detta kan då föra med sig en större transparens, vilket minskar korruption m.m.

Även om det ekonomiska värdet av demokratibefrämjande åtgärder oftast är betydande så är det inte lätt att beräkna.

---

<sup>14</sup> [http://www.elegationen.se/sites/default/files/imce/filer/publikationer/VNR\\_Vagledning\\_i\\_nyttorealiserings\\_1.o\\_med\\_bilagor\\_1.pdf](http://www.elegationen.se/sites/default/files/imce/filer/publikationer/VNR_Vagledning_i_nyttorealiserings_1.o_med_bilagor_1.pdf)

<sup>15</sup> Se t.ex. Open Data Albania <http://open.data.al/cn/>

## RÄTTSSÄKERHET

Transparens och rättssäkerhet går ofta hand i hand. Den myndighetsutövning som syns i maskinläsbara rådata kan enklare granskas och analyseras. De så kallade Stockholmsprinciperna<sup>16</sup>, som förordar att varje stat bör redovisa öppna data om på vilket sätt internetåtkomst begränsas och/eller övervakas, illustrerar hur öppenhet kan stärka rättssäkerheten.

## AKTUALITET OCH SNABBHET

Genom att myndighetsdata exponeras maskinellt kan de distribueras när som helst på dygnet, vilket ger en högre aktualitet än vid traditionell manuell handläggning för utlämnande av information.

Skillnaden mellan att begära ut en handling och att ladda ner och öppna en fil kan vara väldigt stor.

Att få svar på en godtycklig fråga där svaret behöver processas fram i en traditionell databas hos en myndighet kan ta uppåt 6 månader<sup>17</sup>. Tiden för nedladdning av data däremot är ofta mindre än 1 sekund.

En samling dokument (en webbportal) kräver stora insatser för att hållas aktuell. Genom att publicera öppna data och låta olika aktörer själva hämta och bearbeta data kan högra aktualitet uppnås med minskade manuella insatser.

## KVALITET

Antag att ett inmatningsformulär uppmanar användaren att ange kommun. Om ett fritextfält används kan användaren stava kommunens namn fel. Om en användare skriver ”Sundsvall” och en annan ”Sundsvalls kommun” sänker det databasens kvalitet. Om formuläret innehåller en lista med kommuner är risken stor att listan saknar eller har fel information om de kommuner som skapats, bytt namn eller delats efter att formuläret skapades.

Om formuläret däremot regelbundet hämtar en lista med alla Sveriges kommuner från en och samma uppdaterade källa kan inmatning valideras och högsta kvalitet erhållas. På liknande sätt kan många brister i datakvalitet undvikas tack vare öppna data.

---

<sup>16</sup> <http://stockholmprinciples.org/>

<sup>17</sup> <http://3roundstones.com/2012/01/30/publishing-linked-open-data-makes-good-sense-in-tight-economic-times>

Rita en Sverigekarta med ...

Ort/Plats:

Kommun:

Län:

Rita cirklar  km radie

Sök

Exempel: Stockholm, Loshu

Tips: Om du inte är säker på orten kan du använda \* (asterisk) för att ersätta ett eller flera tecken, eller ? för att ersätta ett tecken, t.ex. Göteborg?.

*I detta formulär saknas kommunen Nykvarn, trots att den tillkom för mer än 10 år sedan.*

Högre kvalitet i databaser ger rimligen bättre beslut, färre onödigt skickade brev, färre felaktigt lämnade besked, färre timmar bortkastade på att söka efter nålar i höstackar och så vidare. Hur stora potentiella besparingar detta kan leda till kan delvis kvantifieras genom att analysera vilka kostnader som finns för kvalitetsbrister i befintliga system.

Den transparens som öppna data ger kan tjäna som incitament för kvalitetshöjning. Till exempel hävdar Stockholms handelskammare<sup>18</sup> att den öppna publiceringen av vårdstatistik på Omvård ([www.omvard.se](http://www.omvard.se)) har fått till följd att vårdgivare ökat sina ansträngningar för att framstå i bättre dager i jämförelserna.

## ALTERNATIVKOSTNADER

Om efterfrågan är stor men öppna data saknas, så kan det hända att någon tar fram alternativa öppna data av lägre kvalitet. Väljer användare data av lägre kvalitet än vad som finns så kan det medföra kostnader i samhället som hade kunnat undvikas. Det lär finnas exempel på förlisningar orsakade av navigationssystem för båtar som baserades på bristfälliga öppna sjökortsdata.

## TILLGÅNG TILL INFORMATION

Öppna data kan - särskilt om de återanvänds - öka medborgarnas tillgång till information. Till exempel kan alla möjliga kartor innehålla information om var handikappanpassade toaletter finns, om dessa

<sup>18</sup> <http://www.chamber.se/?id=4371&threadid=37274&p=1&sp=0>

toaleters positioner publiceras som öppna data. Genom att publicera data anpassat för återanvändning snarare än bara en egen trycksak kan data alltså bli mer tillgängliga.

För den som har attraktiva data och vill sprida dem maximalt är det ofta klokare att lägga sina resurser på att publicera ett API eller en rådatafil än att själv utveckla t.ex. en mobilapp eller en snygg webbpresentation. Långt ifrån alla appar och webbpresentationer blir kioskvältare. Det är svårt att lyckas med tajming, formuleringar, formgivning, marknadsföring osv. Därför är det sannolikt att data når fler om många olika entreprenörer och andra mellanhänder får möjlighet att paketera dem på sina egna sätt. Detta gäller förstås bara om det faktiskt finns ett antal mellanhänder som vill återanvända data, men det är inte alltid så lätt att förutsäga vilka data som efterfrågas.

## SÖK- OCH HITTBARHET

Länkade data ger betydligt bättre förutsättningar för att söka, sortera, filtrera och visualisera data än ostrukturerade samlingar av dokument. Facettbaserad sökning är ett bra exempel: användaren kan successivt tillföra sökkriterier i olika dimensioner (t.ex. göra avgränsningar på ålder, pris, kvalitet, fritextmatchning...) och för varje nytt kriterium kan träffbildens reduceras, men även valbara alternativ i övriga dimensioner.

9 entry filtered from 46 originally ([Reset All Filters](#))

sorted by: [entry-type](#) and [labels](#); then by... •  grouped as sorted

### Case study (6)

1. [An Intelligent Search Engine for Online Services for Public Administrations](#) (Case study), by Jesús Fernández Ruiz, Municipality of Zaragoza, Spain  
*Activity area:* public institution and eGovernment  
*Application area of SW technologies:* portal and improved search  
*SW technologies used:* RDF(S) and in-house vocabularies  
*SW technology benefits:* explicit content relationships, identify new relationships, and improved search
2. [Geographic Referencing Framework](#) (Case study), by Catherine Dolbear, Ordnance Survey, United Kingdom  
*Activity area:* public institution, geographic information system, and eGovernment  
*Application area of SW technologies:* data integration  
*SW technologies used:* RDF(S), OWL, and in-house vocabularies  
*SW technology benefits:* share and re-use data and automation
3. [iLaw-Intelligent Legislation Support System](#) (Case study), by Pyung Kim, Seungwoo Lee, Hanmin Jung, Mi-Kyoung Lee, Dong Min Seo, Won-Kyung Sung, Beom-Jong You, Jong-Sub Lee, Tae-Wan Kim, and Seong-Hwan Park, KISTI and MOJ, Korea  
*Activity area:* public institution, legal, and eGovernment  
*Application area of SW technologies:* improved search and data integration  
*SW technologies used:* public datasets, public vocabularies, RDF, SPARQL, OWL, Rules, and in-house vocabularies  
*SW technology benefits:* improved search and dynamic content generation
4. [Integrated, Connected Search Service for Technical Standards Information](#) (Case study), by Hanmin Jung, Pyung Kim, Seungwoo Lee, Mi-Kyoung Lee, Dong Min Seo, Won-Kyung Sung, Beom-Jong You, Dongkwan Jo, Jinyong Jung, and Soohyun Kim, KISTI and KATS, Korea  
*Activity area:* public institution and eGovernment  
*Application area of SW technologies:* improved search and data integration

Search facets:

**Activity area** 1

- 3 automotive
- 2 broadcasting
- 1 cultural heritage
- 1 education
- 9 eGovernment
- 3 energy
- 1 eTourism

**Application area of SW technologies**

- 2 content discovery
- 4 data integration
- 7 improved search
- 5 portal
- 1 semantic annotation

**SW technologies used**

- 8 in-house vocabularies
- 7 OWL
- 1 OWL Lite
- 2 public datasets
- 5 public vocabularies


**Benefits of SW technologies**

*W<sub>3</sub>Cs lista med fallstudier för länkade data ger ett bra exempel på facettbaserad sökning.  
 Prova själv på <http://www.w3.org/2001/sw/sweo/public/UseCases/>*

Observera att sök- och hittbarhet ökar med länkade data, oavsett om de är öppna eller ej. Man kan till och med tänka sig att det finns större nyttor som kan realiseras inom och mellan myndigheter genom användande av (slutna) länkade data än med öppna data.

## MER FULLSTÄNDIG OCH ALLSIDIG INFORMATION

Genom att länkade data använder samma identifierare för samma resurser kan man dra nytta av vad andra har publicerat om de resurser man själv refererar till. Ett bra exempel kommer från webbplatsen Revyu.com, där vem som helst kan recensera vad som helst. En användare, Tom Heath, har givit betyget 5 till filmen Broken flowers. Tack vare att han angivit filmens URL hos DBpedia - [http://dbpedia.org/page/Broken\\_Flowers](http://dbpedia.org/page/Broken_Flowers) - kan Revyu automatiskt förse recensionen med en bild från filmen, namn på regissör osv. Denna information framgår antingen direkt av filmens post i DBpedia (som är en länkad-data-version av Wikipedia) eller av andra data på nätet som också använder samma identifierare.

<p><b>Broken Flowers</b></p> <p><b>Links</b></p> <p>Homepage: <a href="http://www.brokenflowersmovie.com/">http://www.brokenflowersmovie.com/</a>          See Also: <a href="http://en.wikipedia.org/wiki/Broken_flowers">http://en.wikipedia.org/wiki/Broken_flowers</a></p> <p><b>Tags</b></p> <p><a href="#">bill-murray</a> <a href="#">film</a> <a href="#">jessica-lange</a> <a href="#">jim-jarmusch</a> <a href="#">julie-delpy</a> <a href="#">movie</a> <a href="#">sharon-stone</a></p> <p><b>Reviews (1)</b></p> <p>★★★★★ <a href="#">by tom on 30 Jan 2007</a></p> <p>Broken Flowers provides a fantastic vehicle for a classic deadpan Bill Murray performance. The film centers around his character Don, who one day receives a letter from an ex-girlfriend, telling him he has a teenage son. The letter is unsigned, so (with encouragement from his neighbour) he sets off round the country, visiting each the exes who could be the mother of his son. Predictably they're all different in personality and life situation, giving plenty of raw material for awkward silences and dubious encounters. This is great viewing for any Bill Murray fans, or anyone who likes their humour intelligent and a little bit quirky. The soundtrack is also excellent, and deserves a separate review.</p> <p>What do you think of Broken Flowers? <a href="#">Write Your Own Review...</a></p>	<p><b>Broken Flowers</b></p>  <p>directed by <a href="#">Jim Jarmusch</a></p> <p><a href="#">RDF Metadata About Broken Flowers</a></p> <p><a href="#">RDF META</a></p>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

*Automatiskt inhämtad data gör presentationen på Revyu.com mer fullständig*

Även utan data som följer W3Cs principer för länkade data<sup>19</sup> är det förstås möjligt att göra liknande automatiska kompletteringar. Webbplatsen bilpriser.se är ett bra exempel. Givet ett registreringsnummer kan den presentera fakta om bilmodellen, information om ägaren till bilen, prisstatistik, försäkringskostnader, passande reservdelar, länka till liknande bilar på begagnatmarknaden och så vidare. Förutsättningen är att det finns unika identifierare som kan användas vid förfrågningar mot olika databaser. Just när det gäller bilar finns det etablerade sådana identifierare (registreringsnummer, modellbeteckning mm).

<sup>19</sup> <http://www.w3.org/DesignIssues/LinkedData.html>



Värdera din bil direkt med regnr:  Värdera! **GRATIS**

Ny värdering utan regnr »

**VW Sharan 1.8T (150hk) -09**  
48 mån | 7207 mil | Värdeinskningsstartdatum 2008-10-20

Köppris **135 400 kr** Privattpris **112 600 - 123 200 kr**  
Uppskattat pris vid köp hos bilhandlare Uppskattat pris vid köp hos privatperson

Läs mer om bilen Se bilens historik (8)

Basdata / teknisk spec. >

Miljö >

Drivmedel Bensin  
Bränsleförbr. 0,95 l/mil  
CO2 225 g/km

Säkerhet >

Detaljerna med mital, paket etc. för en mer exakt värdering » Visa >

VW Sharan 1.8T (150hk) 2009 Fordon avsett för fler ändamål Manuell **1 418 kr/mån\*** Månadskostnaden baseras på priset: 117 900 kr **Ansök om lån** Swedbank   
\* Eff ränta 0,42%

Liknande bilar till salu Visa alla >

Ove Olofsson Bil AB www.olefssonbil.se **Bilmetro** Bilmetro AB Din bilaffär www.bilmetro.se Svenska Motorkompaniet Vi köper din VW!

BILFÖRSÄKRING BILLÅN HISTORIK DÄCK & FÄLG RESERVDELAR TESTER

Nedan finner du erbjudanden på bildelar till din bil  
På [Skruvat.se](#) kan du hitta ytterligare tusentals bildelar av toppkvalitet till marknadens lägsta priser.

Bränslefilter 143 kr /st  
AUDI A4, A6, CITROËN

Flatblade 99 kr /st  
ALFA ROMEO 33, ALFASUD

Flatblade 99 kr /st  
AUDI A2, CHRYSLER

VÄRDERA LIKANDE BILAR

VW Sharan 1.9 TDI (115hk) 01-10  
VW Sharan 1.9 TDI 4-Motion (115hk) 01-10  
VW Sharan 2.0 (115hk) 01-10  
Visa alla >

*Bilpriser.se sammanställer data från olika källor - vissa är öppna medan andra kostar pengar.*

Med länkade öppna data som följer W3C:s principer blir det enklare och billigare att utveckla denna typ av sammanställningar. De går att skapa snabbt med hjälp av standardkomponenter och kan ofta ske utan att teckna särskilda samarbetsavtal med källan. Idag finns den typen av tjänster för branscher som dels omsätter stora pengar, dels redan har etablerade identifierare för resurser (bilar, flygdestinationer, fastigheter, aktier etc.). Med länkade öppna data sänks trösklarna så att nyttiga sammanställningsfunktioner kan skapas för allt från amatöridrott till tillfälliga forskningsprojekt.

## INNOVATIONER

Tillgång till heltäckande, aktuella data kan ge insikter om besparingsutrymmen och nya affärsmöjligheter. Bara genom att ta del av data kan vi inse vad som kan göras på ett billigare eller effektivare sätt. Överskådliga tabeller, och inte minst visualiseringar av data, kan alltså så frön till innovationer i betraktarens tankar. En del företag tror så mycket på den sortens innovationer att man väljer att öppna upp sina egna data mot kunder och allmänhet. Om kunderna får ta del av vissa interna

data kan de hjälpa oss att bli bättre, resonerar man, enligt konsultfirman Deloitte i rapporten ”Open data. Driving growth, ingenuity and innovation”<sup>20</sup>.

En annan kategori av innovationer exploaterar själva tillgången till data. När det gäller företagsdata är det nystartade engelska företaget Duedil<sup>21</sup> ett tydligt exempel: de sammanställer öppna data om företag och presenterar dessa på ett användbart sätt för den som vill kolla upp (göra due diligence på) en tänkbar affärspartner. I grundutförande är tjänsten gratis men sedan tjänar Duedil pengar på att komplettera med ytterligare data som inte är öppna.

Enligt Deloitte är data den globala ekonomins nya kapital och öppna data blir en allt viktigare del av informationsmarknaden.

## VETENSKAPLIG UTVECKLING

En av de mekanismer som gör att länkade öppna data kan leda till innovationer gör även att länkade öppna data kan leda till vetenskapliga framsteg. När tillgången till data ökar ges också större möjligheter att analysera och dra slutsatser.

Det räcker med att olika aktörer börjar publicera miljörelaterad data på samma sätt för att forskare ska kunna jämföra data från olika håll, se samband och relevanta faktorer och föreslå förbättringar. Det tror iallafall den amerikanska miljömyndigheten EPA som satsar stort<sup>22</sup> på länkade data.

## TEKNISK UTVECKLING

HTML var inte det enda sättet att skapa länkade dokument i början av 1990-talet, och internet var inte det enda globala nätet. En stor del av förklaringen till att just webben slog igenom så snabbt var att alla webbläsare hade en funktion för att visa källkoden. HTML är ett öppet format. Tack vare detta kunde mängder av personer enkelt se hur det gick till att publicera på webben, själva härma detta och på så vis bidra till att webben växte. Programmerare kunde bygga egna webbläsare, egna sökmotorer osv. Detta till skillnad från konkurrerande ”webbar” som t.ex. AOL och CompuServe, där en leverantör ensam kontrollerade både mjukvara och innehåll.

Öppna format, till skillnad från proprietära och dolda format, sänker på så vis trösklarna för deltagande i den tekniska utvecklingen. Länkade data bygger på öppna format. Publicerar myndigheter öppna länkade data så kan andra aktörer kopiera och härma. Alla programmerare kan skriva egna program som drar nytta av dessa data. Öppna format är på det sättet ”generativa” och bidrar till att den tekniska utvecklingen går snabbare och blir mindre beroende av en enskild aktör. Sannolikheten att data går att tolka efter arkivering är också större med standardiserade format och identifierare än med proprietära

---

<sup>20</sup> <http://www.deloitte.com/assets/Dcom-UnitedKingdom/Local%20Assets/Documents/Market%20insights/Deloitte%20Analytics/uk-insights-deloitte-analytics-open-data-june-2012.pdf>

<sup>21</sup> <https://www.duedil.com>

<sup>22</sup> [http://semanticweb.com/the-cpa-explores-a-linked-data-ecosystem\\_b24009](http://semanticweb.com/the-cpa-explores-a-linked-data-ecosystem_b24009)

och skräddarsydda lösningar vilket länge uppmärksammats av svenska arkivmyndigheter som t.ex. Riksarkivet.

## AUTOMATISERING

I viss utsträckning vågar vi idag förlita oss på digitala system för att automatisera viktiga åtgärder. Jag litar till exempel på att min internetbank kommer att betala en viss räkning om jag schemalägger den till ett visst datum, och kanske låter jag en auktionssajt veta mitt högsta bud och för min räkning automatiskt följa med i en budgivning upp till den nivån. Men mitt förtroende för sådan automatisering sträcker sig till händelser inom ett specifikt system som oftast baseras på en robust databas som är enkel att programmera.

Däremot är det oftast inte lämpligt att vidta automatiska åtgärder baserat på information hämtad från webben. Webbsidornas innehåll är alltför svårt att tolka maskinellt för att det ska kunna ligga till grund för automatiserade beslut. Skulle sidans webbredaktör plötsligt göra en designförändring så kanske programmet förväxlar pris med klockslag och gör en usel affär åt mig.

Länkade data (den semantiska webben) ger betydligt större möjligheter för automatisering än de länkade dokument som utgör dagens webb. Om ett program som sköts av en seriös aktör baserar sina åtgärder på data från endast betrodda källor på den semantiska webben så är det rimligt att ge ett sådant program vissa förtroenden.

Faktum är att det finns sofistikerade program (så kallade "reasoning engines") som kan dra slutsatser utifrån data på den semantiska webben. De kan analysera stora mängder data och t.ex. hitta motsägelser även om de inte är uppenbara utan framgår av långa kedjor av data från olika källor. I teorin kan detta leda till att vissa beslut och automatiska åtgärder kan bli bättre underbyggda än de beslut vi idag fattar baserat på en manuell bearbetning av den mängd information vi hinner tillgodogöra oss.

## MYNDIGHETSSPECIFIKA NYTTOR

Det framgår av själva benämningen av myndighetsspecifika nyttor att de kan vara olika och bero på vilken myndighet man talar om. Som tidigare nämnts är det inte heller alltid så att myndighetsnyttan sammanfaller med allmännyttan. Det kan finnas fall där riskerna eller kostnaderna för den enskilda myndigheten med länkade öppna data väger tyngre än myndighetens nytta, men där allmännyttan skulle gynnas mer.

Nedan beskrivs några generella nyttor ur ett myndighetsperspektiv, men varje myndighet bör själv gå igenom vilka nyttor respektive risker och kostnader man skulle få med länkade öppna data.

- **Ökad relevans.** Med ökad användning av myndighetens data kan fler komma att förstå myndighetens roll och uppgifter och därmed få ökad förståelse för myndigheten. En EU-finansierad fallstudie av 21 myndigheter i olika europeiska länder visar också att antalet

återanvändare av data ökade med 1 000 - 10 000 procent när informationen tillgängliggjordes gratis eller till självkostnadspris.<sup>23</sup>

- **Starkare varumärke.** Att tillgängliggöra sin information som öppna och länkade data visar att myndigheten är innovativ, öppen och följer med i utvecklingen, vilket ger en mer positiv bild av myndigheten.
- **Minskad administration.** Genom att kostnadsfritt och automatiskt distribuera data kan en myndighet eliminera behovet att manuellt administrera utlämnande och betalning.
- **Fler samarbetsmöjligheter.** Genom att tillgängliggöra egna data kan en myndighet öppna för nya samarbetsmöjligheter med andra myndigheter eller företag.
- **Nya kunder.** I vissa fall kan öppna data leda till att myndigheten får fler betalande kunder, dvs. att fler efterfrågar delar av myndighetens information eller tjänster som inte släpps gratis. Ett exempel här är kulturinstitutioner, som t.ex. konstmuseer, som genom att tillgängliggöra metadata till reproduktioner av sina verk dels får fler betalande besökare till museet och dels får fler professionella aktörer som är beredda att betala för högupplösta bilder av konstverken.<sup>24</sup>
- **Kunskapsbyggande.** Genom att arbeta med att publicera länkade öppna data (speciellt om man är tidigt ute) skaffar sig myndigheten kunskap och expertis inom området som andra kan behöva och som ev. kan säljas på konsultbasis.
- **Minskad inläsning.** Genom att arbeta med länkade data publicerade på de öppna format som W3C publicerat minskar organisationen sitt beroende av en enskild systemleverantör. Programvaror från olika leverantörer kan tolka och bearbeta data som representeras på öppna format. Detta ger både trygghet och ökade möjligheter till prispress.
- **Ökad datakvalitet** genom möjlighet till öppen granskning. Öppna data kan granskas av många oberoende parter vilket skapar förutsättningar för ökad datakvalitet.
- **Minskade kostnader för egenutvecklade systemlösningar** eftersom informationsöverföring mellan myndigheter och inom den egna myndigheten förenklas om alla parter arbetar med standardiserade och öppna format.
- **Uppnå politiska mål** om förenklat och minskat uppgiftslämnande, tillgänglighet och öppenhet. Öppna data gör det lättare att utbyta och bearbeta information.

Om man ser till myndighetsspecifika nyttor vid ett internt myndighetssamarbete så är det framför allt möjligheten till utökade samarbetsmöjligheter och kunskapsbyggande som bör lyftas fram.

---

<sup>23</sup> POPSIS, Pricing of PSI Study, Deloitte 2011  
[http://ec.europa.eu/information\\_society/policy/psi/docs/pdfs/report/11\\_2012/summary.pdf](http://ec.europa.eu/information_society/policy/psi/docs/pdfs/report/11_2012/summary.pdf)

<sup>24</sup> The Problem of the Yellow Milkmaid. A Business Model Perspective on Open Metadata. Europeana White Paper No. 2, November 2011. <http://pro.europeana.eu/documents/858566/2cbf1f78-e036-4088-af25-94684ff90dc5>

## KATEGORISERING AV NYTTOR

I nedanstående matris, som inspirerats av E-delegationens ”Vägledning till nyttorealiserings”<sup>25</sup> sorterar vi de ovan beskrivna nyttorna med öppna länkade företagsdata utifrån tre dimensioner:

**Färgen** visar hur väl vi kan kvantifiera nyttan, i enlighet med PENG-metodens skala:

- \* **Röda** - kvalitativa nyttor som vi har svårt att värdera kvantitativt
- > **Gula** - kan värderas, men inte exakt
- + **Gröna** - här kan värdet beräknas med någorlunda god precision

**Raden** visar vilken av de tre trenderna den beskrivna nyttan främst härleder från.

**Kolumnen** visar vem som gynnas av nyttan.

Nyttor med öppna företagsdata	För egna myndigheten	För offentlig sektor	För privatpersoner	För företag och org.	För medborgare
Öppna data	* <b>Ökad relevans</b> + <b>Minskad administration</b> > <b>Bättre datakvalitet</b> > <b>Nya kunder</b> * <b>Kunskapsbyggande</b> * <b>Bättre varumärke</b> * <b>Nya samarbetsmöjligheter</b>	> <b>Slipper själv inhämta vissa data</b> > <b>Mer aktuella data</b> > <b>Bättre beslut</b>	* <b>Snabbare tillgång till information</b> * <b>Ökad delaktighet</b> * <b>Större rättssäkerhet</b> * <b>Billigare information</b>	> <b>Billigare informationsförsörjning</b> * <b>Bättre beslutsunderlag</b> * <b>Större rättssäkerhet</b> * <b>Extern produktutveckling</b>	* <b>Ökad insyn</b> * <b>Minskad korruption</b> * <b>Minskat offentligt slöseri</b> * <b>Bättre miljö</b> * <b>Vetenskapliga framsteg</b> * <b>Innovationer</b> * <b>Större rättssäkerhet</b>
Länkade data	* <b>Bättre ordning och reda</b> * <b>Bättre beslutsunderlag</b> * <b>Kan använda standardverktyg</b> * <b>Bättre sökbarhet</b> * <b>Minskat leverantörsberoende</b>	* <b>Bättre förståelse för sammanhang</b> * <b>Bättre sökbarhet</b> * <b>Bättre arkivbeständighet</b>	* <b>Bättre sökbarhet</b>	* <b>Nya affärsmöjligheter</b> * <b>Bättre sökbarhet</b> * <b>Större möjligheter till automatisering</b> * <b>Mindre teknisk inlåsning</b>	* <b>Tekniska framsteg</b> * <b>Vetenskapliga framsteg</b> * <b>Mer fullständig information</b>
Masterdata	> <b>Färre felaktiga utskick</b>	* <b>Snabbare hitta</b> * <b>aktuell</b>	* <b>Bättre service</b>	* <b>Färre onödiga ärenden</b>	

<sup>25</sup> [http://www.elegationen.se/sites/default/files/imce/filer/publikationer/VNR\\_Vagledning\\_i\\_nyttorealiserings\\_i.o\\_med\\_bilagor\\_1.pdf](http://www.elegationen.se/sites/default/files/imce/filer/publikationer/VNR_Vagledning_i_nyttorealiserings_i.o_med_bilagor_1.pdf)

	* Bättre besluts- underlag	information		* Underlättar organisatoriska förändringar	
--	-------------------------------	-------------	--	--------------------------------------------------	--

## EKONOMISKA BERÄKNINGAR AV NYTTAN MED ÖPPNA DATA

En rad positiva beräkningar av den ekonomiska nyttan av öppna data har publicerats under senare år. Inte minst har EU-kommissionen finansierat undersökningar och rapporter som försöker beräkna och beskriva det ekonomiska värdet av att PSI-direktivet implementeras på allt fler områden och i allt fler länder inom EU. Så beräknade t.ex. EU-kommissionen år 2006<sup>26</sup> att det kommersiella värdet av offentlig information på den svenska marknaden uppgick till mellan 226 och 614 miljoner euro. En annan uppskattning från EU-kommissionen från 2011 visar att det övergripande ekonomiska värdet på att öppna myndighetsdata kan uppgå till totalt 40 miljarder euro per år för EU-området, inklusive Norge och Island (EU 27).<sup>27</sup> I samma dokument anges att de totala direkta och indirekta ekonomiska vinsterna från öppna dataapplikationer och -användning i hela EU 27 kan uppgå till i storleksordningen 140 miljarder euro per år. Siffrorna härrör från en rapport där en konsult på uppdrag av Kommissionen granskat tidigare beräkningar och utifrån dessa beräknat ett nytt ekonomiskt värde.<sup>28</sup>

En rapport från den brittiska motsvarigheten till Riksrevisionen (National Audit Office) diskuterar svårigheterna att på ett rimligt sätt uppskatta vinsterna med öppna data.<sup>29</sup> Baserat på samma analys som EU-kommissionen använde 2011 gjorde den brittiska regeringen en uppskattning att det totala ekonomiska värdet för brittiska offentliga sektorn av att öppna myndigheters data till 16 miljarder brittiska pund. Denna beräkning kan dock jämföras med en studie från 2006 där det brittiska konkurrensverket (Office of Fair Trading) tillfrågade mer än 400 offentliga organ för att identifiera intäkter från att tillgängliggöra data öppet och gratis<sup>30</sup>. Här uppskattades värdet till ca 590 miljoner brittiska pund.

Skillnaden mellan dessa uppskattningar är med andra ord betydande och båda studierna har tydliga begränsningar. Top-down metoder, såsom den som används i EU-studien, tenderar att överskatta det ekonomiska värdet av offentliga data. Enligt rapporten från brittiska National Audit Office beror detta bland annat på att man inte räknar med att myndigheter och företag använder rimliga substitut för öppna

<sup>26</sup> Digital Agenda för Sverige, <http://www.regeringen.se/content/1/c6/17/72/56/5a2560cc.pdf> sidan 28

<sup>27</sup> Open data. An Engine for Innovation, Growth and Transparent Governance.

<sup>28</sup> [http://ec.europa.eu/information\\_society/policy/psi/docs/pdfs/opendata2012/open\\_data\\_communication/en.pdf](http://ec.europa.eu/information_society/policy/psi/docs/pdfs/opendata2012/open_data_communication/en.pdf)

<sup>29</sup> Review of recent studies on PSI re-use and related market developments, G. Vickery, August 2011.

[http://www.google.se/url?sa=t&ret=j&q=&esrc=s&source=web&cd=2&sqi=2&ved=oCCoQFjAB&url=http%3A%2F%2Fec.europa.eu%2Finformation\\_society%2Fpolicy%2Fpsi%2Fdocs%2Fpdfs%2Freport%2Ffinal\\_version\\_study\\_psi.docx&ei=H1t2UNisAqOm4gSonoA4&usq=AFQjCNFGTgoQUvCw8EUprj\\_w6fFL6Qb4mA&sig2=PsmcyCpaUCvkqOpyjiR1cw](http://www.google.se/url?sa=t&ret=j&q=&esrc=s&source=web&cd=2&sqi=2&ved=oCCoQFjAB&url=http%3A%2F%2Fec.europa.eu%2Finformation_society%2Fpolicy%2Fpsi%2Fdocs%2Fpdfs%2Freport%2Ffinal_version_study_psi.docx&ei=H1t2UNisAqOm4gSonoA4&usq=AFQjCNFGTgoQUvCw8EUprj_w6fFL6Qb4mA&sig2=PsmcyCpaUCvkqOpyjiR1cw)

<sup>29</sup> Implementing Transparency [http://www.nao.org.uk/publications/1012/implementing\\_transparency.aspx](http://www.nao.org.uk/publications/1012/implementing_transparency.aspx)

<sup>30</sup> The commercial use of public information (CUPI), Office of Fair Trading, December 2006.  
[http://www.of.t.gov.uk/shared\\_of.t/reports/consumer\\_protection/of.t861.pdf](http://www.of.t.gov.uk/shared_of.t/reports/consumer_protection/of.t861.pdf)

data, om informationen inte finns tillgänglig eller är för dyr. Ersättningsdata kan visserligen vara sämre eller dyrare än om myndigheternas data släpps fri, men det ekonomiska värdet av substituten behöver inte vara noll kronor.

Vidare menar National Audit Office att uppskattningen utgår ifrån att den offentliga sektorn i olika länder är relativt likartad, vilket knappast är fallet. Man använder alltså beräkningsmetoder som passar i ett land för att uppskatta värdet av öppna data i andra länder. Å andra sidan tycks rapporten från Office of Fair Trading underskatta det ekonomiska värdet av informationen. Deras undersökning omfattade inte alla myndigheter som kan inneha data av ekonomiskt värde och därmed bör det samlade ekonomiska värdet vara högre än det som rapporten anger. Som framgår av resonemanget ovan är beräkningar baserade på antaganden alltid osäkra och möjliga att ifrågasätta.

Ett alternativt tillvägagångssätt är att genom empiriska studier se hur olika marknader utvecklas till följd av att offentlig information släpps fri. Ett sådant exempel utgörs av en rapport beställd av EU-kommissionen, kallad MICUS-rapporten, som genom enkäter och telefonintervjuer med företag och andra aktörer som återanvänder offentlig information granskat effekten av att öppna upp offentliga data inom det geografiska, meteorologiska och juridiskt administrativa området inom EU.<sup>31</sup>

Studien publicerades i januari 2009 och jämför med situationen 2002, dvs. innan PSI-direktivet antogs. Rapporten slår fast att trots att PSI-direktivet inte hunnit slå igenom fullt ut inom alla tre områdena så är undersökningsresultaten otvetydigt positiva. Effekten sägs vara störst inom sektorn för geografisk information som rapporteras ha vuxit sedan 2002. Inkomsterna för företag som återanvänder geografisk information har stigit och nya användargrupper har tillkommit och erbjuder nya applikationer för geografisk information. Marknaden för juridisk och administrativ information har också vuxit och inkomsterna stigit, men inte i samma utsträckning som inom den geografiska sektorn. Inom den meteorologiska sektorn hittade författarna färre positiva tecken men även här rapporteras det om en växande marknad för meteorologiska data. Samtidigt visar studien att en majoritet av företagen inte kände till PSI-direktivet eller möjligheterna att använda öppna myndighetsdata. Med ökad kunskap om tillgängligheten på öppna data skulle effekten möjligen vara ännu större.

Slutsatsen från detta avsnitt är att det finns skäl att förhålla sig avvaktande inför existerande uppskattningar av det ekonomiska värdet av öppna data. Uppskattningarna kan både överdriva och underskatta värdet. Däremot kan empiriska studier av marknadsförändringar, t.ex. till följd att PSI-direktivet implementeras på allt fler områden, ge värdefull kunskap. Förhoppningsvis kommer fler empiriska studier att genomföras. Med ett bredare empiriskt underlag skulle man kunna göra säkrare ekonomiska uppskattningar av det ekonomiska värdet av öppna data.

## KOSTNADER OCH RISKER

Även om öppna data erbjuder många möjligheter finns utmaningar, kostnader och risker som måste hanteras. Trots att många är överens om att det finns en stor potential till ökad samhällsnytta och

---

<sup>31</sup> Assessment of the Re-use of Public Sector Information (PSI) in the Geographical Information, Meteorological Information and Legal Information Sectors, Micus Management Consulting GmbH (2009)

kostnadsbesparingar med öppna och länkade data, så är denna potential till stor del fortfarande outnyttjad. Utöver kostnader och risker med att öppna data så finns ett antal förutsättningar som måste vara uppfyllda och några utmaningar att hantera. Liksom i det förra avsnittet beskrivs först allmänna risker. I ett senare avsnitt diskuteras olika typer av myndighetsspecifika risker.

## UTMANINGAR

### NYTTAN LIGGER HOS NÅGON ANNAN

Den kanske största utmaningen inom öppna data är att nyttan ofta uppstår hos någon annan än den som sitter på data. För samhället är det naturligtvis önskvärt att en satsning som kostar 1 mkr hos myndighet A genomförs om den ger nyttor värda 10 mkr på nationell nivå, även om nyttan hos myndighet A är mindre än 1 mkr.

### BRIST PÅ EKONOMISK ANALYS

Det är ovanligt att myndigheter analyserar den kommersiella potentialen i sina data och vilken efterfrågan som finns för olika typer av information. Möjligen borde myndigheter uppmanas att prioritera att offentligt tillgängliggöra de datamängder som har störst ekonomisk potential. Men i praktiken kan detta vara komplicerat eftersom det är svårt för myndigheten att veta vilka data som har en sådan potential - en av grundtankarna med öppna data är ju att data kan återanvändas till helt andra ändamål, även syften som den ursprungliga ägaren av data inte kunnat tänka sig. En möjlighet vore att tydligare redovisa vilken information myndigheten har och som man ännu inte delar öppet. Då skulle t.ex. entreprenörer kunna anmäla sitt intresse och eventuellt beskriva sina idéer om hur dessa data skulle kunna bli ekonomiskt värdefulla. Sådana intresseanmälningar skulle kunna utgöra ett underlag för en myndighets prioritering.

### BRIST PÅ INFORMATION

Det är ofta för få som känner till att viss data existerar och finns öppet tillgänglig. Det visar t.ex. MICUS-studien.<sup>32</sup> Där framkom i telefonintervjuer att kunskapen är låg både inom den privata och den offentliga sektorn. Sex av tio organisationer som återanvänder meteorologisk data inte kände till PSI-direktivet. Inom det juridiska området var motsvarande siffra sju av tio och inom geografiska sektorn åtta av tio.

Återanvändarna kände inte heller till vilka rättigheter de har att återanvända data av detta slag. Ökad information riktad till företag och andra aktörer som återanvänder myndighetsdata kan vara ett sätt att öka kunskapen och därmed nyttjandegraden.

---

<sup>32</sup> Assessment of the Re-use of Public Sector Information (PSI) in the Geographical Information, Meteorological Information and Legal Information Sectors, Micus Management Consulting GmbH (2009)



## OTYDLIGT VEM SOM ÄGER INFORMATIONEN

En av huvudpunkterna som diskuterades vid en nyligen timad internationell konferens om öppna data<sup>33</sup>, var att det ofta är otydligt vilken myndighet som ansvarar för vilka data. En sådan brist på tydlighet gör det sannolikt svårare och dyrare för företag och andra aktörer att efterfråga och återanvända informationen.

## BRIST PÅ TYDLIGHET HUR MAN FÅR ÅTERANVÄNDA INFORMATIONEN

I MICUS-studien framkommer att många som återanvänder myndighetsdata uppfattar det som lika viktigt eller viktigare att licensreglerna, dvs. vad man får och inte får göra med data, är klara och tydliga som att priset sänks. Tydliga och långsiktiga spelregler brukar vara nyckelfaktorer för att stimulera företagsamhet.

## DATA ERBJUDS INTE I ÖPPNA FORMAT

En del data görs bara tillgänglig i format som begränsar möjligheterna till återanvändning eller som kräver att den som vill återanvända informationen först bearbetar eller omvandlar den till ett annat format. Det är kostnadskrävande, framför allt om varje återanvändare själv måste göra denna bearbetning. Rent samhällsekonomiskt vore det bättre om den myndighet som äger informationen gör denna bearbetning en gång för alla. Annars uppstår sannolikt en situation där informationen återanvänds i betydligt mindre utsträckning än om den var lätt tillgänglig.

Samtidigt finns det goda skäl att uppmuntra till publicering på vilket som helst format, hellre än att inte publicera alls. Om informationen är efterfrågad finns det säkert någon aktör som är villig att konvertera data till ett öppet format.

## FÖRUTSÄTTNINGAR FÖR PUBLICERING

Det kan finnas många hinder för en myndighet som står i begrepp att exponera information som öppna länkade data. Inte minst handlar det om juridiska konsekvenser, sekretessfrågor och eventuell betalning. Även tekniska och organisatoriska utmaningar måste beaktas.

En aspekt som är central för att få länkade data att bli nyttiga är hantering av unika beständiga elektroniska identifierare. Om ett projekt t.ex. berör domänen ”bilar i Sverige” så är det ganska uppenbart att de registreringsnummer som Transportstyrelsen utfärdar är en bra bas för identifierare. Men för att kunna arbeta med standardteknik för länkade data bör identifierare representeras på URI-form<sup>34</sup>. Detta för att åstadkomma verkligt unika identifierare<sup>35</sup>. Även om svenska registreringsnummer är unika för bilar i Sverige så finns det faktiskt andra länder med snarlika format (t.ex. Finland). Alltså bör vi

---

<sup>33</sup> ”Open Data La Conference: How to create value from Open Data?”, Paris 27 September 2012 <http://www.opendata-laconference.fr/programme.html>

<sup>34</sup> <http://www.ietf.org/rfc/rfc2396.txt>

<sup>35</sup> <http://www.w3.org/TR/webarch/#identification>

i länkade data inte använda formatet XYZ123 utan snarare <http://transportstyrelsen.se/fordon/XYZ123>. Man måste både ta ställning till *vem* som ska utfärda unika identifierare, och *hur* de ska representeras på URI-format.

E-delegationens ”Vägledning för återanvändning av information”<sup>36</sup> besvarar många av de juridiska frågorna och är en bra guide, så vi nöjer oss i övrigt med att hänvisa till det dokumentet. Den brittiska regeringens vitpaper ”Open Data White Paper - Unleashing the Potential” ger också god vägledning.

## FÖRUTSÄTTNINGAR FÖR ANVÄNDNING

Redan idag finns möjlighet att köpa maskinläsbara data eller motsvarande onlinetjänst som kan användas för att validera/slå upp grundläggande företagsinformation. Vid många systemutvecklingsprojekt undersöks dock inte den möjligheten. Det finns helt enkelt ingen budget för att förhindra ”smutsiga” data i databasen, och ansvaret för korrekt grunddata överläts ofta slentrianmässigt på användarna.

Låt säga att företagsdata skulle finnas och vara gratis, men att aktiebolag hanteras av en leverantör, stiftelser av en annan, myndigheter av en tredje, osv. Då skulle de troligen inte heller användas, eftersom det skulle kosta ganska mycket att utveckla kopplingar till alla dessa leverantörer.

Låt säga att en uppslagstjänst finns, att den är gratis och känd av utvecklaren, men att det finns frågetecken kring hur länge den kommer att finnas kvar, eller att det finns frågetecken kring hur beständigt tjänstens API är, eller hur driftsäker tjänsten är, eller hur säker den är, eller hur snabbt den svarar. Eller att det är tekniskt komplicerat att använda sig av den. Då skulle den troligen inte heller användas.

Det finns alltså ganska många faktorer som i praktiken kan utgöra hinder för realisering av de nyttor som är förknippade med tillgång till öppna företagsdata.

## KOSTNADER

Det finns dels direkta kostnader förbundna med att exponera information som öppna länkade data, dels indirekta. De direkta kostnaderna innefattar:

- kostnad för att inskaffa data
- kostnad för att distribuera data
- kostnad för att förvalta data

Kostnaderna i dessa fall kommer sannolikt att variera betydligt eftersom olika data är olika dyra att införskaffa och förvalta. Förvaltningskostnaderna blir högre ju oftare informationen förändras.

Kostnad för att bearbeta data beror på hur data tillgängliggörs. Existerande information kan vara insamlad och lagrad på ett sätt eller i ett format som gör att den först måste bearbetas för att kunna göras

---

<sup>36</sup> <http://feedback.edelegationen.se/vlpsi/>

öppen och möjlig att länka till. Bearbetningen kan exempelvis innebära att data behöver konverteras eller anonymiseras så att enskilda uppgifter inte kan spåras tillbaka till enskilda personer.

Indirekta kostnader för öppna länkade data är bl.a.

- intäktsbortfall - en myndighet som tidigare tagit betalt för viss data förlorar denna intäkt om informationen istället tillgängliggörs gratis. Även här kan kostnaden variera alltifrån ett betydande inkomstbortfall till obetydliga förändringar.
- behov av ombyggnation av webbplats - existerande webbplatser eller andra gränssnitt kan behöva anpassas för att kunna härbärgera och distribuera öppen och länkbar data.

Brittiska riksrevisionen menar att den största kostnaden i samband med öppna data är personalkostnader (t.ex. för att införskaffa, bearbeta och distribuera data). Vanligen saknas metoder för att kunna göra rimliga kostnadsuppskattningar eftersom kostnaderna ofta är utspridda som många små poster på många personer.

## ALLMÄNNA RISKER

Det finns risker, som inte ska underskattas, med att öppna och länka offentlig information. Bland de viktigaste riskerna kan följande nämnas:

- **Hot mot enskildas integritet** till följd av ”oönskad transparens”. Detta kan uppstå genom samkörning av databaser som kan göra kombinerad information känslig, även om beståndsdelarna var och en för sig inte är känsliga. I vissa fall kan denna kombinerade information aktivera sekretesslagar m.m. Därmed kan det finnas större hinder för länkade data än för öppna data.
- **Oklar ansvarsfördelning när unika elektroniska identifierare ska skapas för olika objekt.** Att skapa identifierare för olika ting är grundläggande i arbetet med länkade data. Behovet av att göra kopplingar till informationsobjekt i andra domäner än den egna är vanligt förekommande. Genom att utforma elektroniska identifierare på ett enhetligt sätt där ansvaret för myntandet av dem är tydligt fördelat blir det möjligt att undvika dubbellagring av information. Om ansvarsfördelningen är oklar leder det lätt till att enskilda initiativ skapar egna identifierare för samma sak.
- **Brant inlärningskurva.** Länkade data öppnar en komplicerad värld som kan vara svår att orientera sig i. Den som vill läsa in sig på området stöter snabbt på begrepp som ontologi, schema, URI, RDF, trippellager, artificiell intelligens och så vidare. Förekomsten av alla dessa begrepp gör att det tar tid att förstå och skapar risk för att missförstånd kan uppstå. Dessutom används ofta andra programvaror än de som länge varit vanliga, vilket kräver att teknisk personal kan behöva lära nytt.
- **Risk för feltolkning.** Data är inte alltid enkla att tolka. Ett skäl kan vara att det sammanhang som tidigare omgav informationen kan ha försvunnit i samband med att en myndighet publicerade data öppet. Utan detta sammanhang kan det krävas betydligt större expertkunskaper för att rätt tolka informationen. Än mer komplicerat kan det vara att rätt tolka samkörda eller länkade data

som skapar nya samband och ny information. Feltolkade eller missförstådda data är en risk som blir allvarigare när missförståndet kan spridas till andra. Inom länkade data används ofta ontologier för att tydliggöra vilka begrepp som är utbytbara, motsatta eller relaterade på andra sätt. Ibland saknas dock ontologier. Att ta fram sådana kan i och för sig vara kostsamt och kräva både kommittéarbete och kompromisser, men samtidigt ökar det förståelse och kan förbättra kommunikationen inom berörd bransch.

- **Bristande datakvalitet.** Med ökad användning av data kan kvalitetsbrister få större konsekvenser. Ibland behöver man göra avvägningar mellan datakvalitet och publiceringshastighet. Även om snabb publicering är ett stort värde så kan det ibland vara bättre att publicering av data dröjer till dess kvaliteten har hunnit kontrolleras.
- **Risken för bedrägerier ökar** med ökad information t.ex. om kontraktsdetaljer och betalningsinformation. Bedragare kan försöka använda informationen till att skicka falska fakturor eller omdirigera betalningar. Brittiska erfarenheter visar att omfattande bedrägeriförsök genomfördes framför allt på kommunal nivå, i samband med att offentliga data släpptes fri.<sup>37</sup> Möjligen kan man befara att annan typ av brottslighet också underlättas med ökad tillgång till adressuppgifter, förmögenhetsinformation, geografiska data, kalenderdata m.m.
- **Maktförskjutning.** Välorganiserade, kunniga och välfinansierade aktörer (företag eller organisationer) kan skaffa sig en mycket stark position gentemot enskilda medborgare. Genom att data ibland är svårgenomtränglig för en privatperson eller kräver samkörning med andra data för att bli användbar, kan välorganiserade aktörer med större kunskap vinna mer på öppenheten än privatpersoner. Mycket offentlig data har ju samlats in och använts till medborgarnas bästa. Om den släpps fri för andra aktörer, kan helt andra agendor komma i förgrunden. Därmed kan en ny maktbalans inträda.
- **Förväntade nyttor kan utebli** vilket bl.a. skett i Danmark när staten erbjöd öppna adressdata<sup>38</sup>. Det visade sig att ca 15 procent av alla adresser saknades i den öppna tjänsten, eftersom det var hemligstämplad information. Därmed kunde inte alla klara sig med enbart den öppna informationen och tvingades till andra alternativ.
- **Lägre kvalitet för slutanvändaren.** Genom att data skiljs från den ursprungliga och säkra källan finns risk att data korrumpas av tredje part till men för slutanvändaren. Ett exempel är en läkemedelskatalog, av ungefär samma slag som FASS<sup>39</sup>, som publicerades på ett sätt som låg nära öppna länkade data. Dessa data som var professionellt formulerad information om läkemedel hämtad ur läkemedelskataloger återpublicerades på ett lättillgängligt sätt, men *utan* informationen om biverkningar.

---

<sup>37</sup> Implementing Transparency, [http://www.nao.org.uk/publications/1012/implementing\\_transparency.aspx](http://www.nao.org.uk/publications/1012/implementing_transparency.aspx)

<sup>38</sup> [http://www.adresse-info.dk/Portals/2/Benefit/Value\\_Assessment\\_Danish\\_Address\\_Data\\_UK\\_2010-07-07b.pdf](http://www.adresse-info.dk/Portals/2/Benefit/Value_Assessment_Danish_Address_Data_UK_2010-07-07b.pdf)

<sup>39</sup> <http://www.fass.se/LIF/home/index.jsp>

## MYNDIGHETSSPECIFIKA RISKER

Som framhölls i tidigare avsnitt om myndighetsspecifika nyttor så sammanfaller inte alltid allmänintresset och enskilda myndigheters intressen. Detta kan gälla även för risker med öppna och länkade data. Det faktum att en myndighet förlorar kontrollen över viss data genom att den blir öppen behöver inte vara något problem ur allmännyttans synvinkel, men kan givetvis vara problematiskt för den enskilda myndigheten.

Exemplen på risker som beskrivs nedan bygger alltså på att myndigheter emellanåt utgår från sina egna intressen. Liksom ifråga om nyttorna så behöver varje myndighet göra en riskbedömning ur sitt eget perspektiv.

- **Förlorad kontroll.** Om data släpps fri har myndigheten inte längre kontroll över hur den används. Det kan finnas en risk för dåligt rykte i samband med att återanvänd data tolkats fel eller missförstått och att detta smittar av sig på myndigheten.
- **Försämrat varumärke.** Genom att släppa data fri riskerar myndigheten att förknippas med återanvändare som den inte vill kopplas samman med.
- **Förlust av erkännande eller tillskrivning.** Genom att släppa data fri kan myndigheten komma att mer sällan bli citerad eller tillskriven upphovsrätten. Det kan i sin tur innebära att myndigheten upplevs som mindre relevant och att färre personer förstår dess roll och uppgifter.
- **Förlust av brukare.** Genom att data släpps fri kan slutanvändarna av informationen eventuellt hämta den på annat sätt än genom myndigheten, vilket återigen kan leda till att myndigheten upplevs som mindre relevant och att färre personer förstår dess roll och uppgifter.

Om man betraktar vilka risker som kan uppstå vid ett internt myndighetssamarbete så förefaller inte någon av de ovan beskrivna riskerna som relevanta.

## NÅGRA TÄNKBARA SCENARION

Här presenteras några framtidsbilder. Bilderna ger en översiktlig beskrivning av hur några situationer skulle kunna vara lösta om något eller några år. Avsikten med bilderna är att belysa vilken nytta semantisk teknik kan skapa i situationer när data ändå inte är helt öppen.

Motsvarande tekniska lösningar kan byggas på traditionell teknik men då till betydligt högre kostnader både för utveckling och underhåll. Traditionella lösningar blir också mycket mer låsta och svårare att anpassa till ändrade krav.

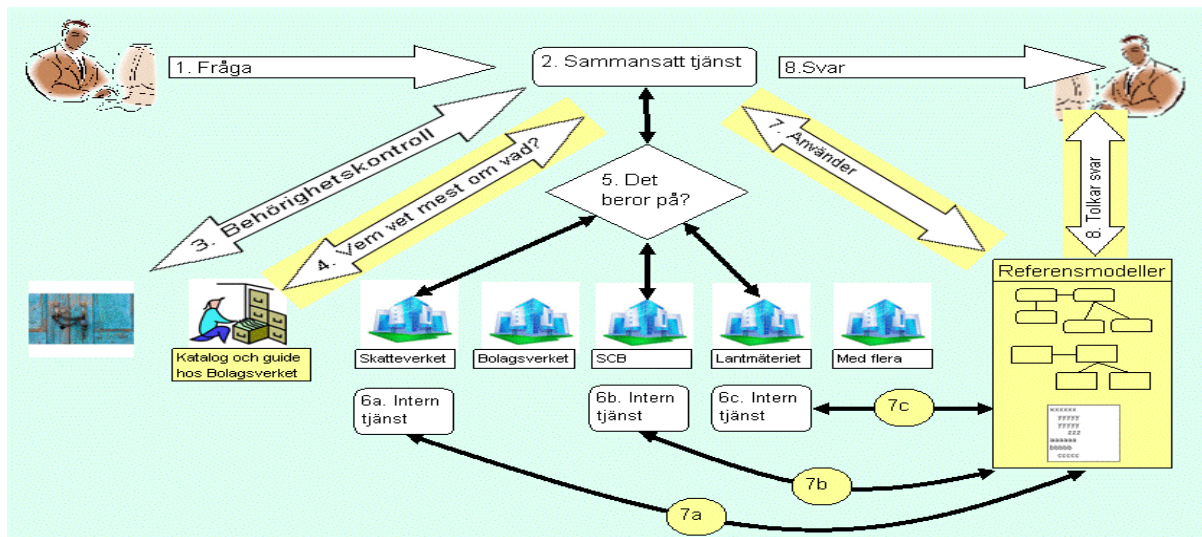
## JAG BEHÖVER GRUNDUPPGIFTER OM EN VISS PERSON.

Någon med rätt att få ta del av grundläggande uppgifter om ett visst företag (fysisk eller juridisk person) ställer en fråga på Internet. Ett samlat svar med de bästa och mest aktuella uppgifter som det offentliga Sverige har ges i XML-format.

Den normala frågeställaren är en myndighet (statlig eller kommunal) som väljer detta alternativ i stället för att själv fråga företagaren (=minskat uppgiftslämnande).

## BESKRIVNING

1. En myndighet skickar in en förfrågan om att få grundläggande uppgifter om en viss samfällighetsförening. Myndigheten legitimerar sig och skickar med någon slags identitet på föreningen.
2. Frågan går till en sammansatt tjänst som leder hela uppgiftsinsamlandet från olika offentliga myndigheter.
3. Tjänsten inleder sitt arbete med att kontrollera om frågeställaren har behörighet att ställa en sådan fråga. En kontroll görs mot en samhällsgemensam funktion för kontroll av behörigheter och rättigheter.
4. Tjänsten skickar en fråga till en gemensam funktion hos Bolagsverket med angivande av företagsform. Funktionen svarar med uppgifter för varje enskild term om var de bästa offentliga uppgifterna finns.
5. Den sammansatta tjänsten fortsätter med att hos aktuella källor efterfråga vad som gäller för just denna bostadsrättsförening. Lantmäteriet är ansvarig för denna typ av juridiska personer så tjänsten börjar med att fråga Lantmäteriet efter huvudmän och kontaktuppgifter. SCB tillfrågas om arbetsställen och Skatteverket om eventuell registrering för moms eller som arbetsgivare.
6. Respektive myndighet har interna tjänster som kan svara på frågorna.
7. Den sammansatta tjänsten använder sig av gemensamt överenskomna modeller för begrepp, informationsobjekt och utbytesformat för att kommunicera med alla inblandade parter. De interna tjänsterna hos respektive myndighet använder sig av samma referensmodeller.
8. Den sammansatta tjänsten svarar med bästa möjliga information i en struktur som bestäms av referensmodellerna. Myndigheten som frågat har inga problem att tolka eller använda svaret då den också har stöd av samma referensmodeller.



## FÖRDELAR MED SEMANTISK TEKNIK

Förutom grundkonstruktionen av den sammansatta tjänsten och de myndighetsinterna tjänsterna krävs mycket lite programmering. Huvuddelen av funktionaliteten ligger i katalogfunktionen (i detta fall hos Bolagsverket) och i användandet av gemensamma referensmodeller.

Den utvecklade tjänsten kan lätt återanvändas i många andra situationer där Bolagsverket vet vem som har den bästa informationen. Det är bara referensmodellerna som behöver anpassas till nya frågeställningar.

Motsvarande katalogfunktioner kan givetvis byggas upp hos andra myndigheter t.ex. hos Lantmäteriet när det gäller geodata eller hos Transportstyrelsen avseende fordon och förare.

## UTMANINGAR

Det krävs en kompetenshöjning inom svensk offentlig förvaltning när det gäller semantiska tekniker.

Sättet att arbeta med verksamhetsutveckling förändras också och blir mer modelldrivet. Det innebär i sin tur att fler resurser krävs från beställaren och mindre från IT-utvecklare.

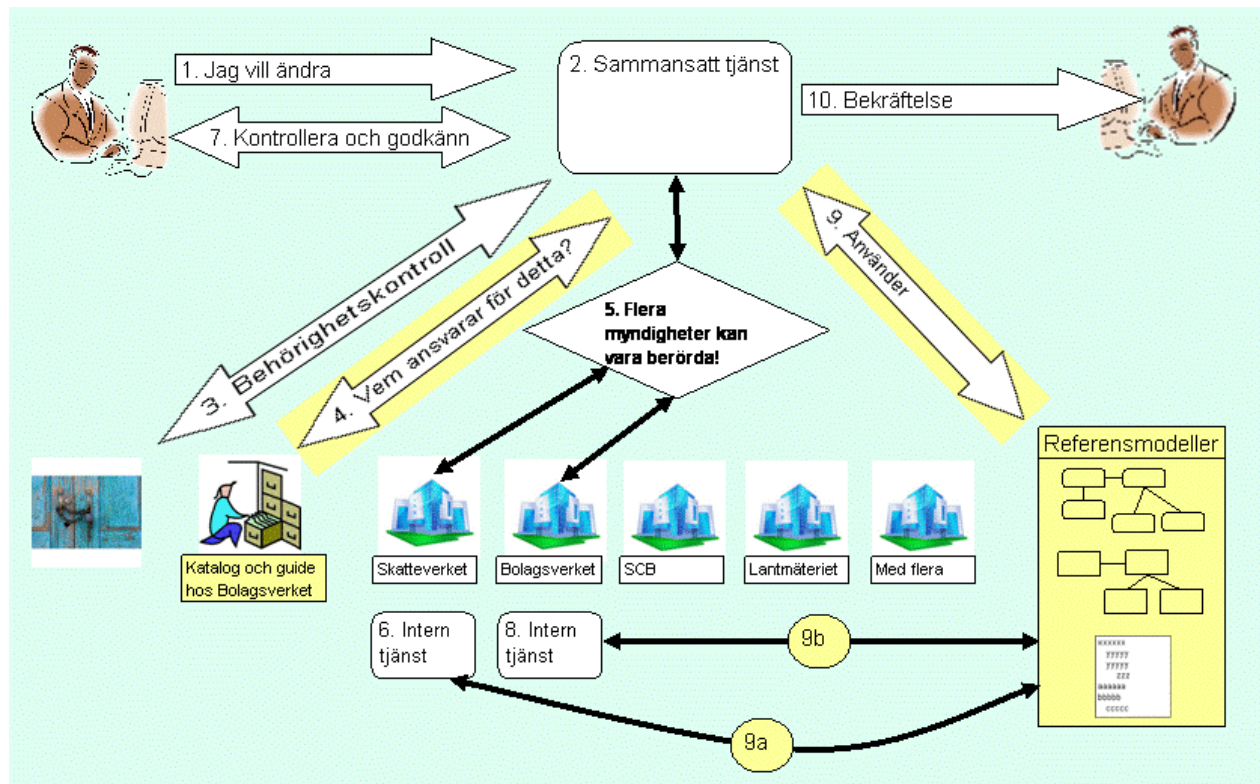
## JAG VILL ÄNDRA MINA UPPGIFTER.

Någon som driver ett företag behöver ändra en uppgift. Det kan handla om att företaget har flyttat eller att någon i styrelsen bytts ut. Det kan givetvis handla även om andra grundläggande uppgifter om företaget. Frågeställaren vänder sig någon myndighet som den ofta har kontakt med eller till t.ex. Verksamt.se. Oavsett kontaktväg möter hen samma dörr – ”Mina uppgifter”

## BESKRIVNING

1. Ett aktiebolag har bytt ut en medlem i styrelsen. En firmatecknare vill göra en uppdatering av de offentliga uppgifterna om bolaget. Hen skickar en begäran om att få uppdatera bolagets uppgifter. Hen kan nå denna tjänst oavsett om hen vänder sig till Bolagsverket, Verksamt.se, Skatteverket eller någon annan offentlig aktör. (=No wrong door). Frågan innehåller också uppgifter om identiteten på den nye styrelsemedlemmen och vem som ersatts samt den roll i styrelsen den nye medlemmen har.
2. Frågan går till en sammansatt tjänst som administrera det offentliga arbetet med uppdateringen.
3. Tjänsten inleder sitt arbete med att kontrollera om uppgiftslämnaren har behörighet uppdatera bolagets uppgifter.
4. Tjänsten skickar en fråga till en gemensam funktion hos Bolagsverket med angivande av företagsform. Funktionen svarar med uppgifter om vilken myndighet som har uppdateringsansvaret för varje enskild term för just denna företagsform. I detta fall Bolagsverket.
5. Flera myndigheter är inblandade eftersom Skatteverket ansvarar för grunduppgifter om fysiska personer.
6. En fråga skickas till Skatteverket som lämnar svar om styrelsemedlemmens namn och adressuppgifter.
7. Svaret presenteras för uppgiftslämnaren som får godkänna detta.
8. Den kompletta anmälan går till Bolagsverket som handlägger ärendet i vanlig ordning.
9. Den sammansatta tjänsten använder sig av gemensamt överenskomna modeller för begrepp, informationsobjekt och utbytesformat för att kommunicera med alla inblandade parter. De interna tjänsterna hos respektive myndighet använder sig av samma referensmodeller.
10. En bekräftelse lämnas till uppgiftslämnaren.





## FÖRDELAR MED SEMANTISK TEKNIK.

Fördelarna är de samma som i det först redovisade scenariot. Genom att det blir möjligt att relativt snabbt utveckla och underhålla lösningar av denna typ kan det ett samhälle där medborgaren har kontroll över sina egna uppgifter byggas.

## UTMANINGAR

Frågan om medborgarens kontroll över information om sig själv kommer att ställas på sin spets.

Publiceringen av offentliga tjänster är också en fråga som måste lösas. Målet är att en användare alltid ska kunna ha lätt att hitta den rätta offentliga tjänsten i den Internet-miljö som man är van att arbeta med.

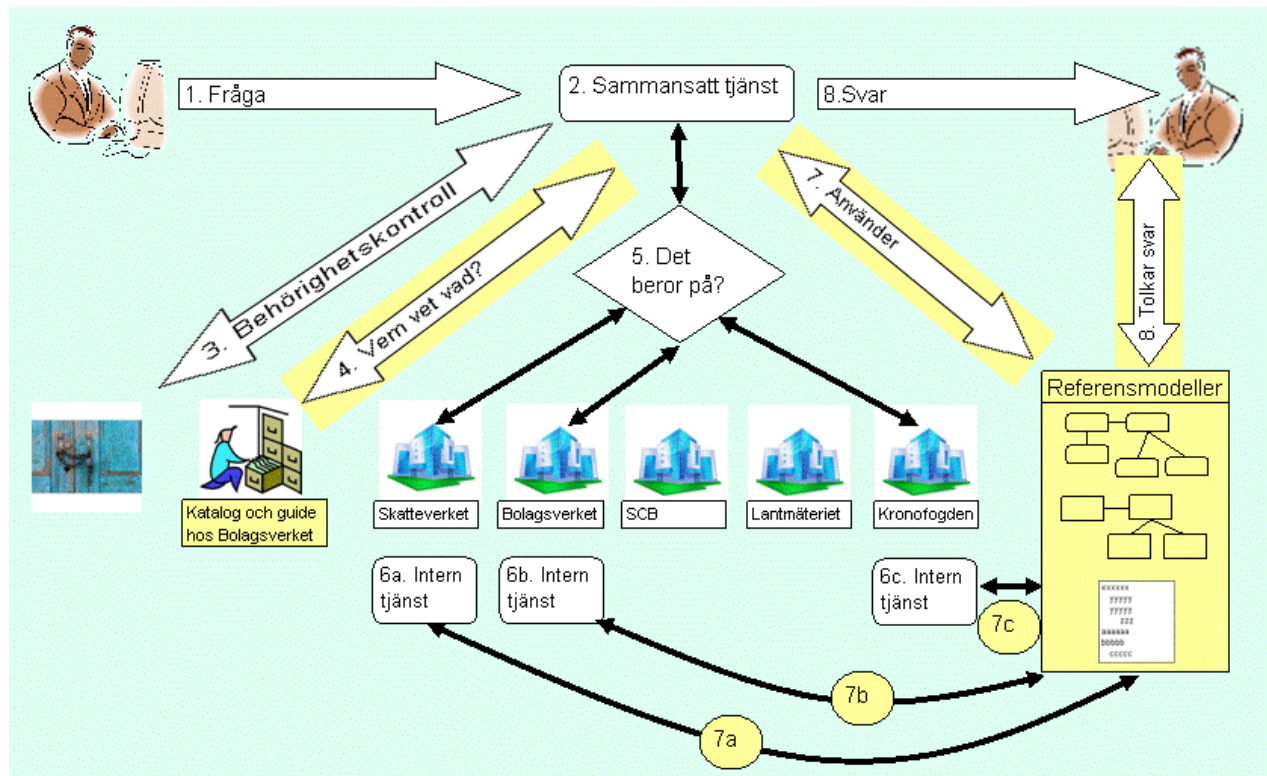
Uttrycket "No wrong door" står för att man alltid ska kunna hitta rätt ingång till kommunikation med det offentliga.

## OFFENTLIG UPPHANDLING

Trafikverket ska genomföra en offentlig upphandling. I stället för att kräva intyg från anbudsgivarna på att de är välskötta företag skickas en fråga till en sammansatt offentlig tjänst. Svaret kommer direkt och är dagsaktuellt.

## BESKRIVNING

1. Trafikverket skickar in en förfrågan om att få uppgifter om en av anbudsgivarna. Myndigheten legitimerar sig och skickar med någon slags identitet på företaget.
2. Frågan går till en sammansatt tjänst som leder hela uppgiftsinsamlandet från olika offentliga myndigheter.
3. Tjänsten inleder sitt arbete med att kontrollera om frågeställaren har behörighet att ställa en sådan fråga. En kontroll görs mot en samhällsgemensam funktion för kontroll av behörigheter och rättigheter.
4. Tjänsten skickar en fråga till en gemensam funktion hos Bolagsverket med angivande av att det handlar om offentlig upphandling. Funktionen svarar med uppgifter för varje enskild term om var de legalt krävda uppgifterna finns.
5. Den sammansatta tjänsten fortsätter med att hos aktuella källor efterfråga vad som gäller för företaget i samband med offentlig upphandling.
6. Respektive myndighet har interna tjänster som kan svara på frågorna. I detta fall tillfrågas Skatteverket, Kronofogden och Bolagsverket.
7. Den sammansatta tjänsten använder sig av gemensamt överenskomna modeller för begrepp, informationsobjekt och utbytesformat för att kommunicera med alla inblandade parter. De interna tjänsterna hos respektive myndighet använder sig av samma referensmodeller.
8. Den sammansatta tjänsten svarar med bästa möjliga information i en struktur som bestäms av referensmodellerna. Trafikverket som frågat har inga problem att tolka eller använda svaret då den också har stöd av samma referensmodeller.



## FÖRDELAR MED SEMANTISK TEKNIK

Det här är ett exempel på hur de fördelar som redovisats i det första fallet kan fungera. Det finns ganska små skillnader mellan en fråga om grunduppgifter om en person och den information som krävs vid en offentlig upphandling. Genom att ändra en parameter i anropet till de myndighetsinterna tjänsterna och byta referensmodell kan denna lösning enkelt åstadkommas.

## UTMANINGAR

Det finns ganska många utmaningar kopplade till frågor om behörighet. Fullmakter kan gälla i vissa fall men inte i andra. En viss officiell roll, t.ex. verkställande direktör kan ha vissa befogenheter i vissa sammanhang men sakna dem i andra. Firmatecknare kan ha en begränsad rätt att agera på egen hand, etc. Dessa frågor bör lösas på nationell nivå.

## SLUTKOMMENTAR

I rapporten beskriver vi ett antal möjligheter och utmaningar med öppna data. Rapporten utgör ett bidrag till projektet "Öppna upp företagsdata!" som drivs under hösten 2012 av ett antal olika aktörer från offentlig och kommersiellt sektor från Norge och Sverige.

Syftet med rapporten är att erbjuda ett underlag för diskussion för personer som ska fatta beslut kring om en myndighet ska börja erbjuda öppna data. Rapporten är på intet vis en komplett introduktion till ämnet öppna data utan mer en förhoppningsvis lättläst introduktion.

Om svensk offentlig sektor samverkar t.ex. vad gäller att ta fram riktlinjer, stöd, katalogtjänster, standarder, vokabulärer och taxonomier så finns goda förutsättningar för att Sverige skulle kunna dra nytta av en utveckling mot mer öppna data.

## BILAGA I: VAD MENAS? - TERMINOLOGISK ORDLISTA MED DEFINITIONER

Som en del i projektet Öppna upp företagsdata har begreppsanalyser gjorts av de centrala begrepp som figurerar i rapporten. Här följer en terminologisk ordlista med de termer och definitioner som projektgruppen står bakom.

---

term	<b>allmän handling</b>
definition	dokument som har inkommit till eller är upprättat hos en statlig eller kommunal myndighet och som förvaras hos myndigheten
anmärkning	De flesta allmänna handlingar är offentligt tillgängliga och benämns då <i>offentlig handling</i> .
källa	Terminologiceentrum TNC: <i>Basord i våra fackspråk</i> (2012).
kommentar	En allmän handling kan vara en text, en bild, en bandinspelning, en film eller annan information som exempelvis har lagrats i en dator eller en mobiltelefon, såsom e-post, SMS-meddelanden, uppgifter i en databas eller en bild av en elektronisk faktura.

---

term	<b>betydelsemässig</b>
synonym	semantisk
definition	som avser betydelsen hos ett språkligt uttryck
anmärkning	I samband med datamängder syftar substantivet <i>semantik</i> ofta på den del av datamängden (t.ex. taggning) som gör att tolkning av betydelsen av övriga data kan göras automatiskt av ett datorprogram.
källa	Språkrådet, Norstedts Akademiska Förlag: <i>Språkriktighetsboken, Termförklaringar</i> (2005), s. 384-404.

---

term	<b>data</b>
definition	representation av fakta, idéer eller liknande i en form lämpad för överföring, tolkning eller bearbetning av människor eller av automatiska hjälpmedel
anmärkning	I strikt mening är det skillnad mellan data och information. Data blir information när någon har tolkat innebörden av data. Många gånger behöver man inte hålla isär

begreppen data och information. Men exempelvis vid överföring mellan datorer eller lagring i datorminnen är det data, inte information, som hanteras.

I många sammansättningar används förledet *data-* om behandling av data med hjälp av dator, som i *datakommunikation*, eller mer allmänt och övergripande, som i *datasystem* och *datakonsult*.

källa Terminologiceentrum TNC: *Basord i våra fackspråk* (2012).

term **delvis strukturerade data** (eng. semi-structured data)

definition data som inte är fullständigt strukturerade men som innehåller någon form av uppdelning i semantiska delar och där hierarkier av poster och fält upprätthålls i datamängden

exempel XML-data, e-postdata

källa Baserad på engelska Wikipedia, webbadress <[http://en.wikipedia.org/wiki/Semi-structured\\_data](http://en.wikipedia.org/wiki/Semi-structured_data)>, hämtad 2012-11-26.

term **företagsdata**

definition data om juridiska personer och om fysiska personer som är enskilda näringsidkare

källa Projektet Öppna upp företagsdata.

term **förädlade data**

definition rådata som bearbetats, sammanställts, kombinerats med andra data eller på annat sätt gjorts lättare att tolka

källa Projektet Öppna upp företagsdata.

term **information**

definition innebörd hos data

anmärkning I strikt mening är det skillnad mellan data och information. Data blir information när någon har tolkat innebörden av dem. Många gånger behöver man inte hålla isär termerna data och information. Men t.ex. vid överföring mellan datorer eller lagring i datorminnen är det data, inte information, som hanteras.

källa Terminologiceentrum TNC: *Basord i våra fackspråk* (2012).

term **länkade data**

definition (på webben:) strukturerade data som följer RDF-regler och som görs tillgängliga via HTTP-protokollet

källa Baserat på definition i *More Structure, More Terminology and (hopefully) More Clarity* (2007), webbadress <<http://www.mkbergman.com/391/more-structure-more-terminology-and-hopefully-more-clarity/>>, hämtat 2012-11-26.

term **länkade öppna data**

definition (på webben:) länkade data som också är öppna data

term **maskinläsbar**

definition som kan läsas i digital form för hantering i en dator

källa Projektet Öppna upp företagsdata.

term **masterdata**

definition datamängd som innehåller en organisation centrala data och som behöver användas av flera system, applikationer och processer

exempel: data om produkter, material, leverantörer, kunder

källa Projektet Öppna upp företagsdata.

term **offentlig handling**

(se allmän handling)

term **offentliga data**

definition data i offentliga handlingar

anmärkning Här avses med offentlig sådant 'som rör eller utförs av stat eller kommun'. *Offentliga data* avser alltså data från statliga eller kommunala myndigheter ("handlingar i det offentliga").

Offentliga data behöver inte vara öppna utan kan vara försedda med olika begränsningar (t.ex. vad gäller vidareutnyttjande, förfoganderätt eller avgifter).

källa Projektet Öppna upp företagsdata.

term **ostrukturerade data**

definition	data som inte har någon känd struktur
exempel	text (som inte är taggad), bilder, video, ljud
källa	Projektet Öppna upp företagsdata.

---

term	<b>proveniens (eng. provenance)</b>
definition	ursprung
anmärkning	En datamängds proveniens är en faktor vid bedömning av dess kvalitet.

---

term	<b>rådata</b>
definition	data som samlats in från en källa och som inte har genomgått bearbetning eller annan manipulation
exempel	obehandlad statistik till en undersökning, data direkt från en fotosensor
källa	Svenska Wikipedia, hämtat 2012-11-26.

---

term	<b>slutna data</b>
definition	data som inte finns allmänt tillgängliga
källa	Projektet Öppna upp företagsdata.

---

term	<b>strukturerade data</b>
definition	data som organiserats på ett sätt som medger tillförlitlig identifiering på typnivå av enskilda faktauppgifter
källa	Baserat på text i Europeiska kommissionen beslut 2011/833/EU, < <a href="http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:330:0039:0042:sv:PDF">http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:330:0039:0042:sv:PDF</a> >.
kommentar	Om data t.ex. innehåller en postadress, ska det entydigt gå att hitta olika delar av adressen (gatuadress, nummer på gata, postnummer, postort etc.).

---

term	<b>tillgängliga data</b>
definition	data som finns allmänt tillgängliga
källa	Projektet Öppna upp företagsdata.

---



term	<b>öppna data</b>
definition	data som finns allmänt tillgängliga via internet och som fritt och utan kostnad får användas, återanvändas och spridas
källa	Projektet Öppna upp företagsdata.

---

term	<b>öppna offentliga data</b>
definition	offentliga data som också är öppna data

---

term	<b>öppen standard</b>
definition	standard vars beskrivning publicerats och är allmänt tillgänglig utan kostnad, som utarbetats och utvecklas i en öppen process av en organisation där alla har rätt att delta och vars immateriella rättigheter – dvs. eventuella patent – görs oåterkalleligt tillgängliga utan royalty
källa	Baserat på text i <i>European Interoperability Framework for Pan-European Egovernment Services</i> , webbadress < <a href="http://ec.europa.eu/idabc/servlets/Docd552.pdf?id=19529">http://ec.europa.eu/idabc/servlets/Docd552.pdf?id=19529</a> >, hämtad 2012-11-26.

---

