



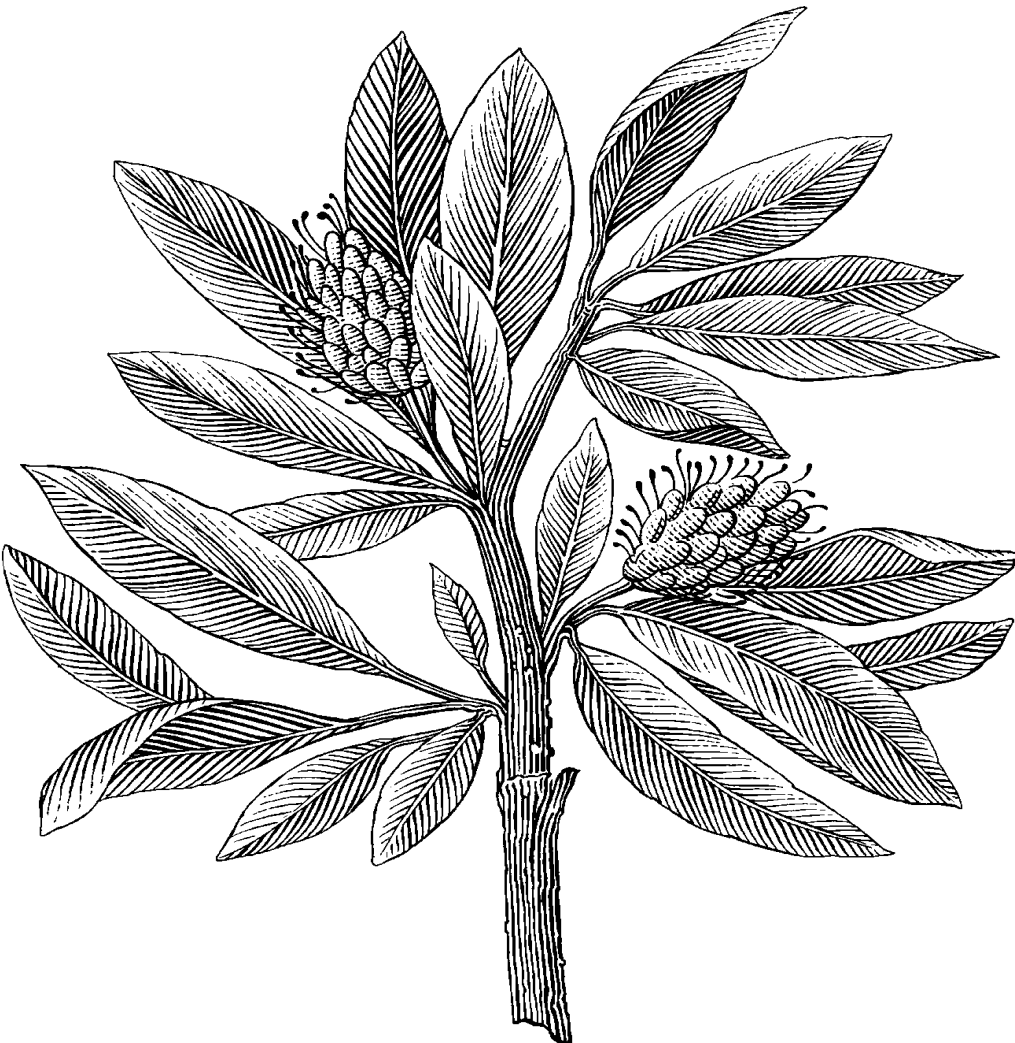
**Linnéuniversitetet**

Kalmar Vaxjö

Datavetenskap, självständigt arbete

## Länkade öppna data

*En studie om fördelar med öppna och länkade data  
för organisationer och dess användare*



*Författare:* Marike Grinde

*Handledare:* John Häggerud

*Examinator:* Johan Leitet

*Termin:* VT2015

*Ämne:* Datavetenskap

*Nivå:* Högskoleexamen

*Kurskod:* 1DV42E

# Abstrakt

Offentlighetsprincipen, som går ut på att alla har rätt att ta del av allmänna handlingar, har funnits i Sverige sedan 1766 och regleras i Sveriges grundlag. I och med dagens informationssamhälle kom 2010-07-01 den så kallade PSI-lagen, med syfte att underlätta enskildas användning av handlingar som tillhandahålls av myndigheter för att främja utvecklingen av en informationsmarknad.

*Öppna data* är en term som används för att beskriva digital information som tillhandahålls fritt och elektroniskt. *Länkad data* är en teknik för detta.

Som står att läsa i denna rapport rekommenderas denna teknik av såväl webbens standardiseringsorgan W3C och webbens grundare Tim Berners-Lee som av den svenska statliga förvaltningsmyndigheten Vinnova, sorterad under Näringsdepartementet. Procentuellt är användandet av tekniken länkad data av svenska myndigheter relativt låg, strax under 5%.

Det här arbetet utreder genom en litteraturstudie vilka som är fördelarna, för en organisation och dess användare, med att tillhandahålla öppna data som länkad data. Resultatet av litteraturstudien används därefter för att i två praktiska exempel utvärdera om de teoretiska fördelarna återspeglats i praktiken. Resultatet påvisar ett stort antal fördelar med länkade öppna data, och att många (men inte alla) av dessa har kunnat återspeglas i praktiken eller förväntas uppnås inom kort.

De slutsatser som uppsatsen kommer fram till är inte överförbara till andra praktiska exempel, men de punkter som identifieras i litteraturstudien kan dock användas som hypoteser att testa på liknande exempel vid fortsatt forskning inom området.

**Nyckelord:** Länkad data, Öppna data, Länkade öppna data, Myndighetsdata

**Tack** till Anna Berggren, Libris och Johan Carlström, Riksantikvarieämbetet som bidrog med värdefull information som gjorde denna studie möjlig. Till John Häggerud, Linnéuniversitetet, för hjälp och vägledning under arbetets gång.

# Innehållsförteckning

## 1 Inledning

### 1.1 Bakgrund

#### 1.1.1 Öppna data

#### 1.1.2 Länkade öppna data

### 1.2 Problembeskrivning och syfte

### 1.3 Frågeställning

#### 1.3.1 Begränsningar

### 1.4 Tidigare forskning

## 2 Teori

### 2.1 Grundtanken med länkad data och den semantiska webben

#### 2.2.1 Möjlighet för maskiner att hitta besläktad information

## 3 Metod och material

### 3.1 Val av metod

#### 3.1.1 Litteraturstudie

#### 3.1.2 Kvalitativ undersökning

##### 3.1.2.1 Intervjuer

### 3.2 Metoddiskussion

#### 3.2.1 Kritik mot litteraturstudien

#### 3.2.2 Kritik mot den kvalitativa undersökningen

## 4 Resultat

### 4.1 Allmän information

### 4.2 Sammanlänkning med andra datakällor

### 4.3 Nya och oväntade användningsområden

### 4.4 Datat ökar i värde

### 4.5 Förbättrad användarupplevelse

### 4.6 Sökmotoroptimering

### 4.7 Bättre struktur och interna fördelar

### 4.8 Återanvändning av data

### 4.9 Semantik

### 4.10 Övriga frågor

#### 4.10.1 Länkade öppna data hos andra myndigheter

#### 4.10.2 Negativa erfarenheter

## 5 Analys och diskussion

### 5.1 Hur kan länkade öppna data förbättra webben för tillhandahållare av digital information?

#### 5.1.1 Vad är de teoretiska fördelarna med länkade öppna data?

#### 5.1.2 Går dessa fördelar att verifiera i praktiska fall där länkade öppna data tillämpats?

### 5.2 Metodreflektion

## 6 Avslutning

### 6.1 Slutsats

## 6.2 Förslag på vidare forskning

### 6.2.1 Hypoteser

#### Referenser

#### Bilagor

Bilaga 1 - Intervju med Anna Berggren, Libris

Bilaga 2, Intervju med Johan Carlström, Riksantikvarieämbetet

# 1 Inledning

Detta kapitel beskriver bakgrunden till ämnesområdet och introducerar läsaren till grundläggande begrepp och valt problemområde. Här presenteras tidigare forskning, uppsatsens syfte och de frågeställningar som uppsatsen ämnar besvara.

## 1.1 Bakgrund

Sedan 1766 har Offentlighetsprincipen funnits i Sverige. Denna innebär att allmänhet och massmedia har rätt att ta del av allmänna handlingar och insyn i statens och kommunernas verksamhet [1].

I och med dagens digitaliserade informationssamhälle trädde 2010-07-01 den så kallade PSI-lagen (2010:566) i kraft, vars syfte är att främja utvecklingen av en informationsmarknad genom att underlätta enskildas användning av handlingar som tillhandahålls av myndigheter [2].

### 1.1.1 Öppna data

*Öppna data* är ett begrepp som använts och tillämpats allt mer de senaste åren. Allt fler företag, kommuner och myndigheter har valt och väljer att öppna upp sin data för allmänheten [3] [4].

För begreppet "Öppna data" finns ännu ingen formell internationellt antagen definition, men definitioner och principer finns formulerade av Open Knowledge Foundation<sup>1</sup> och

Open Government Working Group<sup>2</sup>. Definitionen är "*Open data and content can be freely used, modified, and shared by anyone for any purpose*" och innebär att data kan kallas öppna data om vem som helst fritt och oberoende av syfte får använda, återanvända och distribuera datat med som största motprestation krav att ange källa eller att datas måste delas på samma sätt.

Principer för öppna data som tagits fram och definierats av Open Government Working Group är följande [5]:

1. **Komplett:** Information som inte innehåller personuppgifter eller lyder under sekretess görs tillgänglig i så stor omfattning som möjligt. Detta gäller särskilt databaser med material som skulle kunna vidareförädlas.
2. **Primär:** Information skall så långt det är möjligt tillhandahållas i originalformatet. Bild- och videomaterial skall tillhandahållas i högsta möjliga upplösning för att möjliggöra vidareförädling.
3. **Aktuell:** Information skall tillgängliggöras så snabbt som möjligt så att värdet av den inte försvinner. Det bör finnas mekanismer för att automatiskt kunna få information om uppdateringar.
4. **Tillgänglig:** Information görs tillgänglig för så många användare som möjligt för så många ändamål som möjligt.
5. **Maskinläsbar:** Informationen är strukturerad på ett sätt som möjliggör maskinell bearbetning och samkörning med andra register.
6. **Fri:** Informationen är tillgänglig för alla utan krav på betalning, eller inskränkningar i form av licensvillkor och registreringsförfaranden.

---

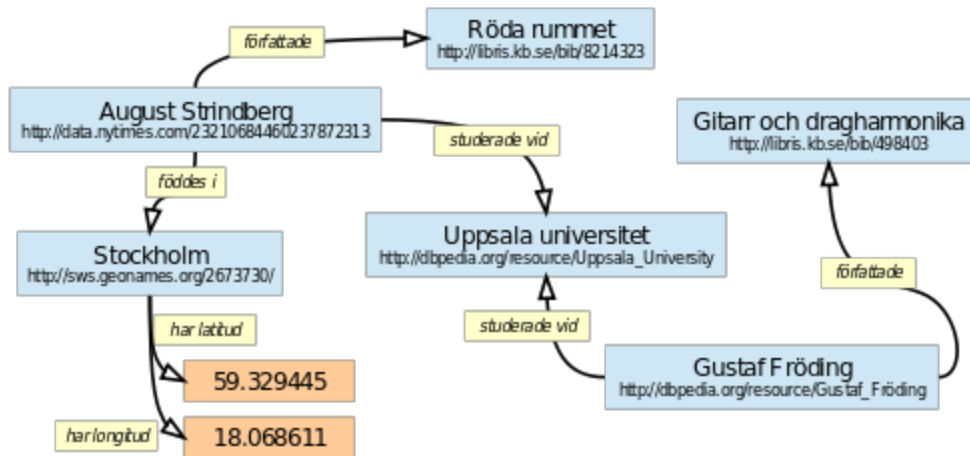
<sup>1</sup> <https://okfn.org/>

<sup>2</sup> <http://opengovdata.org/>

7. **I ett öppet format:** Det format informationen lämnas i följer en öppen standard<sup>3</sup>, alternativt är dokumentationen till formatet fritt tillgänglig och fri från patentlicensvillkor.

### 1.1.2 Länkade öppna data

Länkad data är en metod där man genom att strukturera data med unika identifierare för varje informationsobjekt möjliggör att data och information kan kopplas samman mellan flera datakällor.



**Figur 1.1:** Exempelbild som illustrerar hur länkad data kan fungera, där informationsobjekten har unika identifierare i form av HTTP-URL:er [6].

Länkade öppna data är således öppna data som publicerats i det länkade dataformatet beskrivet i exempelbilden ovan.

För att uppmuntra ägare av data - statliga och privata - till att publicera sin öppna data som länkade öppna data, presenterades 2010 en femstjärnig bedömningsmodell [7], där stjärnor erhålls beroende på hur väl data tillhandahålls online. Bedömningsmodellen är publicerad av webbens grundare Tim Berners-Lee<sup>4</sup> och standardiseringsorganet World Wide Consortium (W3C).

Detta är bedömningsmodellen:

1. Den första stjärnan får man för att tillhandahålla data med en licens som möjliggör fri användning.
2. Den andra stjärnan om ovanstående krav uppfylls, samt om informationen är i ett maskinläsbart format.
3. Den tredje stjärnan om ovanstående krav uppfylls, samt om data är publicerad i en öppen standard<sup>5</sup>.

<sup>3</sup> [http://sv.wikipedia.org/wiki/%C3%96ppen\\_standard](http://sv.wikipedia.org/wiki/%C3%96ppen_standard)

<sup>4</sup> [http://sv.wikipedia.org/wiki/Tim\\_Berners-Lee](http://sv.wikipedia.org/wiki/Tim_Berners-Lee)

<sup>5</sup> [http://sv.wikipedia.org/wiki/%C3%96ppen\\_standard](http://sv.wikipedia.org/wiki/%C3%96ppen_standard)

4. Den fjärde stjärnan om ovanstående krav uppfylls, samt om det är länkad data i rekommenderade W3C-standarder (till exempel i RDF och SPARQL), med unika identifierare.
5. Den femte stjärnan om ovanstående krav uppfylls, och om datat dessutom är sammanlänkat med andra dataset.

## 1.2 Problembeskrivning och syfte

Öppna data som länkad data är något som rekommenderas och uppmuntras av W3C, vars huvudsakliga syfte är att säkerställa webbens långsiktiga tillväxt. Den femstjärniga utvärderingsmodellen som anges i föregående kapitel presenterades för 5 år sedan, vilket är en lång tid i webbutvecklersammanhang. Fortfarande är dock länkad data något som inte tillämpas i stor utsträckning. Av totalt 102 st publicerade datakällor inom svensk offentlig sektor är i dagsläget 2 st femstjärniga och 3 st fyrstjärniga [8]. Det innebär en totalt användning på strax under 5%.

Syftet med denna rapport är att utreda vilka fördelar länkade öppna data medför, samt att försöka utvärdera om dessa fördelar kunnat återses i praktiken. Förhoppningen är att rapporten ska besvara om ett utökat användande av länkad data i samband med öppna data kan föra webbens utveckling framåt.

## 1.3 Frågeställning

Denna uppsats har som avsikt att besvara följande frågeställningar:

- Hur kan länkade öppna data förbättra webben för tillhandahållare av digital information?
  - Vad är de teoretiska fördelarna med länkade öppna data?
  - Går dessa fördelar att verifiera i praktiska fall där länkade öppna data tillämpats?

### 1.3.1 Begränsningar

Det finns länkad data som inte uppfyller kraven för öppna data, exempelvis i form av licenser och krav på hur datat får användas, det vill säga länkad data som inte uppfyller den första stjärnan i bedömningsmodellen som nämns i kapitel 1.1.2. Denna uppsats behandlar inte länkad data som inte uppfyller detta grundkrav för öppna data.

Dessutom begränsar sig denna uppsats till att bara titta på länkade öppna data på en nationell nivå. Anmärkas bör även att inte *alla* praktiska fall där länkade öppna data använts granskas, utan endast två. De fall som valts att utvärderas är de två och enda femstjärniga datakällorna av de som finns listade på opengov.se [5], där man publicerar öppna datakällor inom den svenska offentliga sektorn.

## 1.4 Tidigare forskning

I rapporten *Länkad data som ett alternativ för att publicera öppna offentliga data* [9] undersöks hur svenska kommuner idag arbetar med att publicera information och hur länkad data skulle kunna underlätta utnyttjandet av det publicerade öppna datat ytterligare. Slutsatsen i rapporten är att kommunerna saknar och behöver standardiserade tekniker för att framtidssäkra, underlätta och effektivisera sitt arbete med öppna data. Man menar att länkad data är en lösning för detta, då det finns tydliga W3C-standarder för detta format. Man skriver också att den öppna datan ökar i värde om den blir länkad.

Den här uppsatsen undersöker inte huruvida länkad data som standardiserad teknik kan underlätta för offentlig sektor som vill publicera öppna data.

Däremot finns den andra slutsatsen från rapporten, att öppna data som länkad medför en värdeökning, med som underlag i litteraturstudien (kapitel 2) och som ett av underlagen som testas i den kvalitativa undersökningen (kapitel 4).



## 2 Teori

Detta kapitel presenterar det teoretiska förarbete som gjorts kring länkade öppna data, och som ligger till grund för undersökningen i kapitel 4.

Här tittas på vad som är grundtanken med länkad data och därefter avhandlas vilka fördelar länkad data förväntas ge.

### 2.1 Grundtanken med länkad data och den semantiska webben

Den “semantiska webben” är ett begrepp myntat av Tim Berners-Lee. Begreppet handlar om att genom olika tekniker göra inte bara datat i sig maskinläsbart, utan även innebörden i datat - *semantiken* i den information som finns tillgänglig på webben. Tanken med semantiska webben är att gå från en webb byggd av dokument till en webb byggd av data på samma sätt som i databaser. Målet med detta är att möjliggöra för datorer att göra ett nyttigare arbete och att utveckla system som stödjer pålitliga interaktioner över nätverket [10]. För att detta ska kunna vara möjligt behöver en stor mängd av datat vara publicerat i samma typ av maskinläsbara format. Informationen behöver även vara publicerad på så vis att man kan utläsa dess relationer. Det är denna samling av relationer och inbördes datamängder som kallas länkad data, och data publicerat på detta vis är helt grundläggande för en semantisk webb [11].

Tim Berners-Lee kom på idén med länkade öppna data i början av 2000-talet i samband med att han blev frustrerad av att så mycket information är extremt svåråtkomlig då den ligger skyddad i olika databaser, bakom inloggnings, lösenord och inuti olika system som är svåra att lära sig. Han såg en potential i all dold information, om den skulle vara tillgänglig för alla i sitt råa format. Han menar att om all data och information är sammanlänkad kan det underlätta exempelvis för forskare som behöver mycket information inom ett specifikt ämne. Länkad data är teknologin för en sådan sammanlänkning. Idén går ut på att sätta unika HTTP-identifikatorer till varje objekt istället för - som tidigare - enbart till hela dokumentet. I samband med sammanlänkning av objekt ska man även tala om vilken relation objekten har, för semantiken [12].

Fyra grundregler finns formulerade för hur länkad data bör publiceras (översatt) [13]:

1. Använd URI:er för att namnge och identifiera objekt.
2. Använd HTTP-URI:er så att referenser kan skapas och människor kan slå upp objekten.
3. Tillhandahåll maskinläsbar information om objekten när dessa URI:er anropas. Informationen ska vara i standardiserade format som t.ex RDF<sup>6</sup> och SPARQL<sup>7</sup>.
4. Inkludera länkar till andra URI:er, så att nätverket av länkad data utvidgas och berikas.

---

<sup>6</sup> [http://sv.wikipedia.org/wiki/Resource\\_Description\\_Framework](http://sv.wikipedia.org/wiki/Resource_Description_Framework)

<sup>7</sup> <http://en.wikipedia.org/wiki/SPARQL>

## 2.2 Fördelar med länkade öppna data

### 2.2.1 Möjlighet för maskiner att hitta besläktad information

W3C skriver i sina standarder om Semantiska webben att när flera dataset länkas samman och externa länkar tillhandahålls, möjliggör man för program att utnyttja information från andra dataset som annars inte skulle vara åtkomlig [10].

Detta argument, att länkad data möjliggör både för människor och maskiner att lätt hitta besläktad information, är något som även nämns i artikeln "Linked data" av Tim Berners-Lee [13].

Också i det Vinnova<sup>8</sup>-finansierade projektet "Länkade öppna data i Sverige - Portal och nationell statistik"

<sup>9</sup> listar man i en informationstext ökad datainteroperabilitet som den högsta fördelen med länkad data.

Ökad datainteroperabilitet innebär bättre förmåga för olika system och datamängder att kommunicera och utbyta information. Man framhåller också att "formatet möjliggör automation av dataåtkomst och användning" [14].

I den vitbok som tagits fram i samband med den Vinnova-finansierade projektet "Kompetensförstärkning kring länkade öppna data – dialog, webinarier och vitbok" framhålls det som en av de största styrkorna med de standarder som länkad data innebär att det möjliggör för maskiner att lättare kunna utbyta och samköra data. I andra fall kan det vara vanligt att en given standard endast täcker in en del av de behov man har, och att man tvingas komplettera med byggstenar från andra standarder. Med länkad data tillhandahålls en gemensam bas med en mängd standardiserade byggstenar som kan kombineras på olika sätt. Detta underlättar då mottagarna av informationen är okända och många till antalet [15].

W3C menar att integrerad fakta från flera olika informationskällor ger en mycket bättre användarupplevelse. De framhåller också att det gör verktygen kraftfullare och informationen lättare för människor att förstå [11].

Liknande argument framhålls också av en artikel framtagen av projektet "Länkade öppna data i Sverige", man skriver att det innebär en fördel för konsumenten att kunna upptäcka ny intressant data och kombinationer av data [16].

### 2.2.2 Värdet på informationen ökar

R. Carlsson och M. Svensson skriver i sin forskningsrapport att de i sin studie funnit att det reella värdet på öppna data först förväntas uppstå när kvaliteten på datan uppnår länkad data [9]. Att tekniken länkad data ökar värdet på informationen är något som också hävdas av W3C. Tim Berners-Lee skriver i sin artikel "Linked data" att på traditionella HTTP-webbsidor avgörs värdet på informationen inte bara av datat i sig utan även av hur väl det länkar till annan information, och på samma sätt fungerar den semantiska webben, att informationen ökar i värde om den länkas samman med annan information [11].

I "Vitbok länkade öppna data" förklaras lite närmre varför länkad data förväntas öka värdet på informationen: "Det är ofta så att länkar till andra datakällor också ökar förtroendet för data. Dels genom att länkar skapar ett större sammanhang som bekräftar att man hittat rätt data. Dessutom visar existensen av länkar att man ansträngt sig för att placera in sin data i ett större sammanhang och minimera

---

<sup>8</sup> <http://www.vinnova.se/sv/Om-VINNOVA/>

<sup>9</sup> <http://www.ida.liu.se/~evabl45/lodsweden.sv.shtml>

redundans. Detta gör det mindre sannolikt att någon annan redan har eller kommer ta fram en bättre datakälla som gör den aktuella datakällan överflödig” [15].

### 2.2.3 Data kan återanvändas på fler och oväntade sätt

Tim Berners-Lee skriver i artikeln “Linked data” att man genom att *inte* följa hans fyra regler (i denna uppsats nämnda i kapitel 2.1) för länkad data går miste om möjligheten att återanvända sin data på oväntade sätt, och att det är denna oväntade återanvändning av information som är av värde. I samma artikel står också att SPARQL-frågespråket gör det möjligt att ställa frågor baserade på information från flera dataset [11]. Detta står också att läsa i en text från projektet “Länkade öppna data i Sverige - Portal och nationell statistik”, där man listar de främsta mervärden som skapas genom att publicera länkade data, och man skriver att länkad data ger “data som är förberett för uppgifter och användningsområden ej påtänkta vid publiceringen” [14].

På samma projekts portalplats står att läsa att en av fördelarna med länkade öppna data är att datan blir förberedd att använda på fler användningsområden än den tilltänkta. Nya och oväntade användningsområden kan hittas och förverkligas av konsumenterna. Man framhåller också att dess format möjliggör för den som bygger vidare på informationen får automatisk tolkning av datats betydelse [16].

### 2.2.4 Sökmotoroptimering

I “Vitbok länkade öppna data” framhålls förbättrad sökmotoroptimering<sup>10</sup> som en av fördelarna med öppna data som länkad data. Man beskriver att det skett ett paradigmskifte i hur moderna sömaskiner fungerar och att man i större utsträckning söker efter kunskap istället för webbsidor. Grunden för detta är att sökmotorerna numera kompletterar sin indexering av webbsidor med en kunskapsbas. Exempel på sådana är Googles “Knowledge Graph” som introducerades 2012 och Microsofts “Bing Satori Knowledge Base” från 2013. Tekniken länkad data utgör en stor grund för sådana kunskapsbaser. Dels är de kompletterade med existerande “länkade data-vänliga” kunskapsbaser som exempelvis Freebase, Wikipedia, CIA-factbook. Googles “Knowledge Graph” klarar även att läsa in länkade data som JSON-LD eller inbäddat i webbsidor som RDFa [15].

### 2.2.5 Bättre struktur på informationen

På projektet “Länkade öppna data i Sverige - Portal och nationell statistik”s portalplats beskriver man att tekniken länkad data gör att “dina data blir enklare att upptäcka och hämta utan extra arbetsinsats för den som publicerat data”. Man menar att detta är en fördel för tillhandahållaren av datat [16].

I artikeln “Linked data” av Tim Berners-Lee skrivs att eftersom SPARQL-frågespråket gör det möjligt att ställa frågor kombinerade från flera dataset underlättar tekniken länkade data i de fall man har väldigt mycket information att hålla reda på [11].

### 2.2.7 Information kan återanvändas istället för kopieras

På projektportalen för “Länkade öppna data i Sverige - Portal och nationell statistik” listas bland fördelarna med länkade öppna data att det möjliggör för datakonsumenten att länka till och återanvända delar av informationen i en datamängd [16].

---

<sup>10</sup> <http://sv.wikipedia.org/wiki/S%C3%B6kmotoroptimering>

Tim Berners-Lee tar i sitt tal “Open, Linked Data for a Global Community”, upp att länkningen av information gör att man kan genom den här tekniken när man pekar på något vara säker på att man menar samma sak [17].

I “Vitbok länkade öppna data” skriver man “för att ta ett exempel kan data om staden Stockholm se olika ut om det är SJ, SMHI eller Historiska muséet som tillhandahåller den. Samtidigt kan vi tänka oss att en turist kan dra nytta av att få information om Stockholm från dessa tre källor presenterade tillsammans. Ett bra sätt att åstadkomma detta - då det inte alltid går att hitta globala identifierare - är att skapa relationer mellan de olika datauttrycken med hjälp av länkar. En speciell form av länk (owl:sameAs) uttrycker att det det är samma ting (staden Stockholm) som har fått olika identifierare (URI:er).” I samma vitbok förklarar man också att tekniken med att återanvända istället för att kopiera även är till fördel i form av arbetsbesparingar för organisationen som tillhandahåller datat, då många datauttryck blir återanvändningsbara [15].

## 3 Metod och material

I följande kapitel redogörs vilka metoder som används för att besvara uppsatsens frågeställningar. Här motiveras och sammanfattas tillvägagångssättet och genomförandet. Kapitlet avslutas med en metoddiskussion som diskuterar arbetets validitet och reliabilitet.

### 3.1 Val av metod

Eftersom den här uppsatsens frågeställning innehöll två delfrågor där den andra frågan var beroende av den första, delades metoden upp i två delar. Den första delfrågan använde sig av en litteraturstudie, medan den andra frågan använde sig av en kvalitativ undersökning där de svar som identifierats i litteraturstudien granskades närmre genom intervjuer.

#### 3.1.1 Litteraturstudie

Till att börja med behövde frågan besvaras vad som i teorin ansågs vara de största fördelarna med länkade öppna data. Att detta rekommenderas av webbans industrikonstium har beskrivits i rapportens inledande kapitel, men vilka argument ligger till grund för detta? I litteraturstudien utvärderades vilka argument som ligger till grund för rekommendationen och vilka fördelar man menar att länkade data innebär.

I litteraturstudien valdes att i första hand gå till källan - W3C - för att identifiera de mest argumenterade fördelarna. Av samma anledning (det vill säga att komma så nära källan som möjligt till dem som förespråkar länkade öppna data) valdes att också titta på två Vinnova<sup>11</sup>-finansierade projekt, *Länkade data i Sverige - Portal och nationell statistik*<sup>12</sup> och *Kompetensförstärkning kring länkade öppna data - dialog, webinarier och vitbok*<sup>13</sup>. Även den forskningsrapport[9] som nämns i inledningskapitlet användes i litteraturstudien.

Syftet med litteraturstudien var att identifiera ett antal gemensamma argument förespråkade hos ovan angivna källor att senare använda som granskningsunderlag i den kvalitativa undersökningen.

#### 3.1.2 Kvalitativ undersökning

För vidare analys av argumenten som identifierades i litteraturstudien valdes en kvalitativ undersökning som forskningsmetod. Meningen var att i praktiska exempel titta på om de fördelar som förespråkas i teorin fungerat på samma vis i praktiken.

Som nämns i kapitel 1.2 finns i dagsläget två datakällor inom den svenska offentliga sektorn där datat länkats samman med andra datakällor. De två datakällorna är Libris (Kungliga biblioteket)<sup>14</sup> samt K-samsök (Riksantikvarieämbetet)<sup>15</sup> och det är dessa två källor som använts i den kvalitativa undersökningen.

---

<sup>11</sup> <http://www.vinnova.se/sv/>

<sup>12</sup> <http://www.linked-data.se/wiki/Huvudsida>

<sup>13</sup> <http://metasolutions.se/projects/kompetensforstarkning-kring-lankade-oppna-data/>

<sup>14</sup> <http://libris.kb.se/>

<sup>15</sup> <http://www.ksamsok.se/>

Metoden var intervjuer med representanter för de två datakällorna som ansågs ha goda kunskaper om deras länkade öppna data. Den kvalitativa undersökningen syftade framförallt till att ge en fördjupad analys av dessa två enskilda fall.

Svaren från intervjuerna är subjektiva och har inte kunnat generaliseras till andra datakällor med länkade data. De slutsatser som dragits kan däremot användas som hypoteser att testa på liknande exempel vid fortsatt forskning inom området.

#### 3.1.2.1 Intervjuer

På grund av de geografiska avstånden stod valet av intervjuform mellan telefon/video och mail, och metoden som valdes blev mail. Anledningen till detta var dels önskemål om att få svara via mail från en av de intervjuade som hade svårt att få loss tid, men också att frågorna var av sådan karaktär att man ville ge de intervjuade tid att tänka igenom, formulera och möjlighet att kontrollera sina svar.

Intervjuerna genomfördes med sakkunniga hos Kungliga bibliotekets databas *Libris* och Riksantikvarieämbetets länkade data från kulturarvsområdet *K-samsök*. Organisationerna fick själva välja lämpliga personer. Från Libris intervjuades Anna Berggren, projektledare på avdelningen Informationssystem, härnäst kallad N1. Från K-Samsök intervjuades Johan Carlström, förvaltningsledare för K-samsök, härnäst kallad N2.

De frågor som användes i intervjun baserar sig på den litteraturstudien i kapitel 2. De teoretiska fördelar som identifierats i litteraturstudien är de fördelar som informanterna fått bedömda om de praktiskt kunnat sätta i samband med sin länkade öppna data. De första frågorna är dock allmänna informationsfrågor som inte baserar sig på litteraturstudien. Syftet med dessa frågor är att ge en bättre helhetsbild av hur länge tjänsterna funnits och hur sammanlänkningen av information ser ut.

##### Fråga 1:

När publicerades er länkade öppna data?

##### Fråga 2:

Hade ni dessförinnan någon annan typ av öppna data?

##### Fråga 3:

Med vilken/vilka andra datakällor är ni sammanlänkade, och hur använder ni den informationen från dessa datakällor i er tjänst?

##### Fråga 4:

Använder dessa datakällor information från er i sina tjänster?

##### Fråga 5:

Har tekniken länkad data gjort att er öppna data kunnat användas på nya, oväntade vis?  
(Exempelvis applikationer som byggts eller nya dataset som länkat samman med er data)  
Om så är fallet, ge gärna exempel.

Fråga 6:

Upplever ni att sammanlänkningen av er data med andra dataset gjort att värdet på er data ökat?

Fråga 7:

Upplever ni att tekniken länkad data gjort eran tjänst (för användarna) -

- kraftfullare?
- lättare att förstå?
- i stort till en bättre användarupplevelse?

Beskriv gärna hur responsen från användarna varit?

Fråga 8:

Upplever ni att teknologin länkad data förbättrat er synlighet hos sökmotorer?

Fråga 9:

Har teknologin länkad data gett er några fördelar internt i organisationen?

(Exempelvis bättre struktur på informationen)

Fråga 10:

Har tekniken länkad data möjliggjort att ni kan återanvända information istället för att kopiera den?

Om så är fallet, upplever ni att det underlättat er tjänst och/eller arbetet för er i fråga om:

- Att undvika missförstånd (det vill säga att det blir tydligt att det "pekas på" samma sak)
- Arbets- och tidsbesparingar?

Fråga 11:

En fördel som beskrivs med länkad data är att dess standarder möjliggör för maskiner att förstå innebörden av informationen. Har ni i er tjänst dragit några fördelar av detta?

Om så är fallet, beskriv gärna hur.

Fråga 12:

Skulle det för er innebära någon skillnad om fler myndigheter publicerade sin data som länkad?

(Exempelvis att ni isåfall vill bygga vidare på er tjänst ytterligare)

Fråga 13:

Finns något negativt att säga om tekniken länkad data?

(Exempelvis att det inneburit problem för er eller för era användare)

Fråga 14:

Något övrigt ni vill tillägga om era erfarenheter kring länkad data som inte tagits upp i frågorna?

## 3.2 Metoddiskussion

### 3.2.1 Kritik mot litteraturstudien

Kritiskt i litteraturstudien är att två av de tre använda källorna (W3C och Vinnova projekten *Länkade data i Sverige - Portal och nationell statistik* och *Kompetensförstärkning kring länkade öppna data - dialog, webinarier och vitbok*) kan misstänkas ha en partiskt positiv inställning till länkade data, då de ju förespråkar och jobbar för ett utökat användande av detta. I den här har dock detta inte bedömts som något negativt, då det för den här delen av frågeställningen inte fanns något behov av opartiska källor i fråga om att identifiera positiva argument för länkade data.

Litteraturstudien syftar till att kartlägga hur länkad öppna data i teorin kan utveckla och förbättra webben, och det var med ett sådant syfte naturligt att gå till de största förespråkarna för detta. Det kritiska granskandet kommer istället i den kvalitativa undersökningen, där de teoretiska argumenten granskats i konkreta exempel. I litteraturstudien används också den i kapitel 1.4 nämnda forskningsrapport[6] som ett komplement till de två partiska källorna. Detta för att ge litteraturstudien ytterligare en dimension med en vetenskaplig opartisk källa.

### 3.2.2 Kritik mot den kvalitativa undersökningen

I den kvalitativa undersökningen kan till viss del reliabiliteten ifrågasättas då det ju är fysiska personer som intervjuas och således har svaren blivit subjektiva. Däremot kommer svaren att i största möjliga mån analyseras utifrån ett vetenskapligt och opartiskt förhållningssätt i slutdiskussionen (kapitel 5).

En annan risk i den kvalitativa undersökningen är den begränsade validiteten då endast två källor har undersöks. Det har varit en svårighet att hitta praktiska fall där länkade öppna data har tillämpats. Som nämns tidigare i detta kapitel samt i kapitel 1.2 finns inom den svenska offentliga sektorn i dagsläget endast två datakällor som använder sig av länkade data. Ambitionen var att i den här uppsatsen undersöka båda dessa datakällor, vilket har gjorts. Detta ger det en bra täckning av den lilla information som i dagsläget finns att tillgå. Det bör också understrykas att den begränsade marknaden och bristen på källor gör den information som fått fram till extra värdefull.



## 4 Resultat

I detta kapitel presenteras det empiriska material som är resultatet av den kvalitativa undersökningen genomförd via intervjuer beskrivet i kapitel 3.1.2.1. Resultatet är en sammanställning av intervjuerna och kompletteras med citat från de tillfrågade. Intervjuerna finns i sin helhet tillgängliga som bilaga 1 och bilaga 2 längst bak i denna uppsats.

### 4.1 Allmän information

Libris katalog publicerades som länkade öppna data 2008. De hade dessförinnan ingen annan typ av länkade öppna data eller annan typ av öppna data. Ungefär samma tid, mellan 2008 och 2010, drev Riksantikvarieämbetet utvecklingsprojektet K-samsök, ett regeringsuppdrag med syfte att skapa en tjänst som hämtar data från olika minnesinstitutioner. Riksantikvarieämbetet tillgängliggjorde detta i RDF/XML-format via ett API. All data mappades mot ett gemensamt protokoll, vilket skapade möjligheter att koppla ihop data och samsöka. De första dataseten som använde länkade öppna data publicerades 2009. Efter detta har ytterligare data tillkommit och idag är storleken cirka 6 miljoner objekt från ett 50-tal institutioner. Innan projektet erbjöds datauttag om fornminnen i form av shape- och tabfiler.


### 4.2 Sammanlänkning med andra datakällor

Libris publicerade sitt katalogdata som länkade öppna data redan för sju år sedan (2008), men det är först nu de lagrar allt sitt data i grunden som länkad. De har systemarkitekturen för detta färdig, men den har ännu inte satts i drift. Delar av de nya tjänsterna kommer att rullas ut efter hand och där kommer man att använda sig av externa datakällor som t.ex Dbpedia och Geodata. Förhoppningen och tron är att detta kommer att berika både deras katalogdata och söktjänster. Riksantikvarieämbetets data är i huvudsak sammanlänkat internt med sina egna dataset inom K-samsök, men har också länkar till litteratur från Libris, artiklar i Wikipedia och bilder i Wikipedia Commons. Denna data finns sökbar via ett API och används t.ex av söktjänsten kringla. Ett exempel är

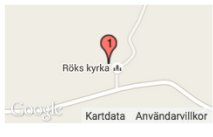
<http://www.kringla.nu/kringla/objekt?referens=raa/fmi/10048200010001>

## Rökstenen, Runristning

Kulturlämning




OBS! Bild från relaterat objekt, se länk t.h.




Kartdata Användarvillkor

### BILDER



### platsr BERÄTTELSE



Rökstenen  
Av Lotte

4 1 2

### ARTIKLAR I WIKIPEDIA

Artiklar i Wikipedia som innehåller information om detta objekt eller närliggande ämnen.

#### ARTIKELN BESKRIVER OBJEKTET


Rökstenen  
Länkad av : Wikipedia

#### ARTIKELN BESKRIVER NÄRLIGGANDE ÄMNE

Runor  
Länkad av : Johan

### BILDER I WIKIMEDIA COMMONS

Bilder i Wikimedia Commons som avbildar detta objekt.



### RELATERADE OBJEKT

#### AVBILDAS AV

Rökstenen  
Typ: Foto  
Källa: Riksantikvarieämbetet  
Länkad av: Automatisk länk

Rökstenen  
Typ: Foto  
Källa: Riksantikvarieämbetet  
Länkad av: Automatisk länk

Rökstenen  
Typ: Foto  
Källa: Riksantikvarieämbetet  
Länkad av: Automatisk länk

[Visa alla](#)

#### EUROPEANA

#### BESKRIVS AV (2)

Der Runestein von Rök in Östergötland  
Typ: TEXT  
Källa: Murberget Länsmuseum Västernorrland  
Länkad av: carwash

Den Runestein von Rök in Östergötland  
Typ: TEXT  
Källa: Murberget Länsmuseum Västernorrland  
Länkad av: Johan

#### LITTERATUR FRÅN LIBRIS

#### BESKRIVS AV (113)

Titel: Rökstenen  
Författare: Friesen, Otto von, 1870-1942  
Länkad av: Johan

Titel: Inbjudningskrift till ähörande af den offentliga föreläsning, med hvilken professorn i zoologi Axel Wirén tillträder sitt embele  
Författare: Schöck, Henrik, 1855-1947

|             |  |
|-------------|--|
| Typ         | Kulturlämning  |
| Plats       | Län: Östergötland, Kommun: Ödeshög, Landskap: Östergötland, Socken: Rök  |
| Titel       | Rökstenen  |
| Beskrivning | Runsten, ljus finkornig granit, 2,6 m hög, 1,35 m bred (N-S) och 0,25-0,4 m tjock. Stenen är närmast rektangulär, med övre kortsidan ca 30° sluttande mot N, vilket gör att stenen upptill är något spetsig. Runinskrift finns på alla fem sidorna. Stenen står under ett skyddshus som uppfördes år 1933.<br><br>Redan genom sin utformning är inskriften svårtolkad i sin helhet. Den består av en berättande och g... |
| Orientering | 3 m S om kyrkogårdsmur och 30 m V om infartsväg (P-plats).   |
| Terräng     | Plan lermark intill kyrkomur.  |
| Referenser  | Gustavson, H, Rökstenen, 1991 -- Jansson, S B F, Runinskrifter i Sverige, 1984, s. 33 ff. - Brate, E, Östergötlands runinskrifter, Öp 426, s. 224 ff., samt där  |

**Figur 4.1:** Exempelbild från söktjänsten Kringla som visar information om Rökstenen i Östergötland. Datat som visas är Riksantikvarieämbetets information om rökstenen tillsammans med kartinformation från Google, platsberättelser från platsr.se, artiklar beskrivande objektet och närliggande ämnen från Wikipedia, bilder från Wikipedia Commons samt litteratur om objektet från Libris.

På samma vis använder sig Wikipedia och Wikipedia Commons av de persistenta länkarna i K-samsök för att länka mellan en artikel om ett objekt med Riksantikvarieämbetets beskrivning. Exempel:

<http://sv.wikipedia.org/wiki/R%C3%B6kstenen>

### 4.3 Nya och oväntade användningsområden

Libris har ännu inte kunnat se om tekniken länkad öppna data möjliggjort att deras öppna data kunnat användas på nya och oväntade vis. Man menar att detta fortfarande är för tidigt att säga, men att förhoppningen förstås är att detta kommer att ske. Riksantikvarieämbetets fornminnesdata har kunnat återanvändas och i flertalet externa applikationer<sup>16</sup>, dock tillägger N2 att “inga är särskilt oväntade”.

### 4.4 Datat ökar i värde

N2 beskriver en värdeökning av datat då “informationen sätts i ett sammanhang på ett annat sätt än tidigare vilket ökar förståelsen och nyfikna användare har möjlighet att klicka sig vidare”. N1 har ännu

<sup>16</sup> <http://www.ksamsok.se/goda-exempel>

inte sett någon värdeökning, men tillägger också att det fortfarande är för tidigt att säga och att “förhoppningen är förstås att det kommer att ske”.

## 4.5 Förbättrad användarupplevelse

På samma vis som med värdeökningen anser N1 att det länkade öppna datat lett till en bättre användarupplevelse då förståelsen och nyfikenheten ökar då informationen sätts i ett sammanhang på ett annat sätt än tidigare. N2 anser det ännu för tidigt för att kunna svara om det länkade datat lett till en förbättrad användarupplevelse.

## 4.6 Sökmotoroptimering

Ingen av informanterna har sett om tekniken länkad data lett till en bättre sökmotoroptimering. N1 svarar att man ännu inte vet och N2 svarar att det är svårt att säga, men “troligtvis inte”.

## 4.7 Bättre struktur och interna fördelar

På frågan om länkade öppna data lett till bättre struktur på den interna informationen eller andra interna fördelar svarar N1: “Inte ännu. I vår organisation har vi stora datamängder som är beskrivna i flera olika format. Vi har utvecklat en systemarkitektur som är formatoberoende, men de olika formaten måste fortfarande förstå varandra för att det ska gå att använda olika datakällor och på så sätt förbättra våra tjänster, såväl externa som interna. Inom vårt team jobbar vi i första hand med nationella tjänster och vårt första mål är att översätta ett biblioteksspecifikt format till standardiserade beskrivningsformer. När detta arbete är avslutat kan vi fortsätta att mappa även andra format för att på så sätt kunna få olika datakällor att berika varandra”. N2 beskriver hur arbetet med länkade öppna data gjort att de ändrat perspektiv och nu ser informationen i större sammanhang i förhållande till andra nationella och internationella datamängder. Länkade data har även gjort dem mer medvetna om standarder och masterdata.

## 4.8 Återanvändning av data

På frågan om tekniken länkad data möjliggjort att man kunnat återanvända information istället för att kopiera den, svarar N1: “Inte ännu. Målet är att vi ska undvika redundans i såväl data som i arbetsflöden. Målet är att effektivisera våra arbetsprocesser och dessutom förbättra datakvaliteten i våra tjänster”. N2 beskriver att man till exempel använder länkar från bilders metadata beskrivande fornlämningar för att länka in bilder i söktjänsten Fornsök.

## 4.9 Semantik

Gällande om länkade öppna data lett till en ökad semantik, har ingen av informanterna sett någon konkret fördel med detta. N1 beskriver att de fortfarande mappar sitt data från det biblioteksspecifika formatet som beskrivet i idag, men säger att man genom att använda standardiserade beskrivningsformer kommer att tillgängliggöra biblioteksdata för internet och maskin.

## 4.10 Övriga frågor

### 4.10.1 Länkade öppna data hos andra myndigheter

På frågan om det skulle innebära någon skillnad om fler myndigheter publicerade sin data som länkad är svaret entydigt Ja.

N1 beskriver att de “skulle kunna använda auktoritetsdata som andra myndigheter är specialister på och därmed förbättra datakvalitén och effektivisera vårt arbete.”

N2 svarar att exempelvis Riksarkivets, Lantmäteriets och Länsstyrelsens data för dem skulle vara “både intressant och nödvändig att kunna länka till och återanvända”.

### 4.10.2 Negativa erfarenheter

Gällande informanternas negativa erfarenheter med formatet länkad data svarar N2 “Tekniken är förhållandevis ny och används inte i så stor utsträckning inom offentlig sektor. Därmed svårt att jämföra och se det långsiktiga i att förvalta denna typ av tekniker”. N1 har ännu inte sett något negativt, men tillägger att det heller inte ska vara någon övertro på det länkade formatet och att “det måste användas rätt för att få effekt”.

## 5 Analys och diskussion

Detta kapitel presenterar en analys av det insamlade resultatet. Kapitlet avser att besvara rapportens frågeställning samt reflektera över resultatet. Sist i kapitlet finns en metodreflektion, där uppsatsens metodval diskuteras.

### 5.1 Hur kan länkade öppna data förbättra webben för tillhandahållare av digital information?

#### 5.1.1 Vad är de teoretiska fördelarna med länkade öppna data?

Litteraturstudien visar på ett stort antal fördelar med öppna data som länkad data. För en tillhandahållare av data verkar fördelarna stora med bland annat ökat värde på datat, bättre struktur samt återanvändning och sammanlänkning istället för kopiering, och ökad sökoptimering. Också för användarna innebär den länkade tekniken många fördelar såsom en bättre upplevelse då det skapar möjligheter till mycket information från många källor utan större ansträngningar som fler klick eller ytterligare sökningar. En annan, mer hypotetisk, fördel för användarna är om webbutvecklarna kan använda det länkade formatet på den öppna datan till att skapa nya intressanta applikationer. Detta skapar fördelar både för användarna, som får ta del av applikationerna, men också för tillhandahållarna, som på så sätt får en större spridning av sin information.

Med så många fördelar kan det tyckas märkligt att så få som endast två myndigheter som i dagsläget använder det länkade formatet. Gissningsvis beror det sparsamma användandet på andra faktorer än det länkade formatet i sig. Myndigheter sitter på enorma mängder data och information. Johan Carlström beskriver exempelvis att Riksantikvarieämbetet för närvarande har cirka 6 miljoner objekt, från ett femtiotal olika institutioner. Att publicera sådana mängder data som öppen och länkad kräver stora resurser såväl ekonomiskt som kunskapsmässigt och juridiskt.

Anmärkningsvärt är även att av de totalt 102 offentliga myndigheter som tillhandahåller data (nämns i inledningskapitlet) [5] är det 74 st som inte uppnått någon stjärna alls. Det vill säga drygt 72% av dem som öppnat upp sin data uppfyller inte grundkravet för *öppna data*, att datat är fritt att använda utan inskränkningar. Detta kan eventuellt tyda på att man överhuvudtaget inom svensk offentlig sektor är dålig på att följa de riktlinjer som finns i samband med publicering av data.

#### 5.1.2 Går dessa fördelar att verifiera i praktiska fall där länkade öppna data tillämpats?

I den kvalitativa undersökning som via intervjuer genomfördes var målet att finna svar på frågan om de fördelar som länkade öppna data påstods medföra gick att verifiera även i praktiska exempel. Med stöd från intervjuerna skulle svaret på den frågan kunna vara både Ja, Nej, och Inte än. Tyngst här torde dock svaret Inte än väga. Man kan se en stor nytta med länkade öppna data då det används fullt ut, som i exemplet med Rökstenen (figur 4.1), där sökobjektet sätts i ett mycket större sammanhang än normalt då man kombinerar fakta från den egna datakällan med externa artiklar, bilder, kartor och litteratur. Detta ger mycket information samlad på samma ställe, och även för användaren information som denne eventuellt aldrig skulle ha kommit i kontakt med om den ej varit sammanlänkad. För tillhandahållaren innebär detta fördelar både i och med att deras tjänster torde bli mer välbesökta och attraktiva, men också ger det en

ökad spridning av, och kanske även ett ökat intresse för, deras information. Som beskrivs i litteraturstudien borde detta också leda till en värdeökning av datat i sig, då det används mer, oftare och då informationen - då den sätts i större sammanhang - löper en mindre risk att bli överflödig. Det är dock svårt att konkret mäta och värdera en sådan värdeökning i pengar.

Libris har bara haft en del av sin data länkad sedan 2008, men kommer snart att publicera all sin data i det länkade formatet från grunden. Detta borde leda till många fördelar då man konkret kommer att kunna se att deras data och information sätts i större sammanhang, på samma vis som man gjort hos Riksantikvarieämbetet. Man har också hos båda informanterna sett eller tror sig kunna se en ökning av återanvändning och minskning av redundans, vilket ju från början var hela grundtanken med länkad data. Ett sådan vinst är förstås av nytta både för organisationer internt men också för informationssamhället i sin helhet.

Båda informanterna beskriver att de själva skulle bli hjälpta om fler myndigheter övergick till det länkade formatet, och att de på så vis både skulle kunna förbättra sin information och effektivisera sitt arbete. I detta faktum borde en stor del av svaret på frågan om de teoretiska fördelarna kunnat verifieras i verkligheten ligga. Det vill säga att även om fördelar till viss del kan skönjas redan nu, borde de största fördelarna uppnås först när det länkade formatet används fullt ut och varit i drift några år, så att det hunnits länkas samman med andra datakällor och satts i nya - väntade och oväntade - större sammanhang.

Webbutveckling går fort, men att strukturera och organisera stora mängder data går långsamt. Så också att ändra organisationers och myndigheters sätt att arbeta. Som nämns i metod-kapitlet har Vinnova drivit två projekt för att främja användandet av länkade öppna data. Förhoppningsvis kommer allt fler myndigheter att övergå till länkade öppna data, när man efterhand utvärderat och sett fördelarna där det används. Det finns dock här en moment 22-risk, då de konkreta fördelarna ej kan ses förrän fler praktisk börjat. Det krävs alltså även ett risktagande från dem som är först ut.

Den litteraturstudie som gjorts i denna uppsats och de argument till fördel för formatet länkad data som där tagits fram, borde förslagsvis kunna användas för ytterligare utvärderingar av datakällor som använder sig av det länkade formatet. I nästa kapitel listas ett antal hypoteser som är resultatet av denna uppsats, och som förslagsvis kan användas som underlag för framtida forskning i liknande undersökningar inom området.

## 5.2 Metodreflektion

Övergripande har metodstrukturen för denna här uppsatsen följt ett tydligt mönster. Det fanns ett tydligt mål för den teoretiska litteraturstudien och vad denna skulle undersöka och presentera. Utifrån detta resultat (presenterat i kapitel 2.2) formulerades ett antal intervjufrågor, och samma intervjufrågor användes i två kvalitativa undersökningar genom intervjuer via mail. Motiveringen till val av mail istället för telefon kändes vid tidpunkten före intervjuerna vettig, men nu i efterhand hade nog inspelade telefonintervjuer hellre varit att föredra. Vissa av svaren var lite otydliga och skulle ha behövt följdfrågor och förtydliganden. Det finns också en risk att informanterna missförstått någon av frågorna de fått, då inte heller dem haft möjlighet att ställa följdfrågor och be om förtydliganden. Dock har det varit intressant

att ställa identiska frågor till de båda informanterna, och det faktum att frågorna varit identiska har gjort det lättare att analysera svaren. Mailformen har också gjort att man minimerat risken att följdfrågorna blivit ledande eller att informanterna inte kunnat förbereda sig på dem och noga fundera igenom och formulera sina svar.

## 6 Avslutning

I detta avslutande kapitlet presenteras uppsatsens slutsats, baserad på den teoretiska litteraturstudien och efterföljande kvalitativa undersökning. Sist i kapitlet ges förslag på vidare forskning inom ämnet.

### 6.1 Slutsats

Litteraturstudien identifierar många fördelar med länkade öppna data, både för dem som tillhandahåller informationen och för deras användare, och därmed indirekt för tillhandahållarna igen. Många av de fördelar som identifierats finns argumenterade på flera ställen och hos flera källor.

Resultatet av undersökningarna stöttar till viss del dessa teorier, men inte alla fördelar har kunnat verifieras. Undersökningarna visar även att den källa som tidigt publicerade all sin data som länkad idag ser fler fördelar än den andra källan, som förväntar ett positivt resultat allteftersom den länkade datan alltmer rullas ut. Båda informanterna uttrycker en önskan om att fler myndigheter ska övergå till det länkade öppna formatet, då detta skulle förbättra datakvalitet, tjänster och även kunna effektivisera arbetet.

Omfattningen på undersökningen gör det svårt att dra stora och generella slutsatser kring huvudfrågeställningen, men resultatet av undersökningen indikerar på att ju mer man använder det länkade formatet fullt ut och över tid, och ju fler som använder formatet och länkar samman, desto fler fördelar för såväl tillhandahållare som slutanvändare.

### 6.2 Förslag på vidare forskning

Detta arbete har riktat in sig på två specifika datakällor där ett antal punkter testats och utvärderats. De punkter som testats kan vara intressanta att testa på fler datakällor, och man kan på så sätt nå ett bredare resultat för större och mer generella slutsatser än dem som dragits i denna rapport.

Hypoteserna som följer kan användas som vägledning vid fortsatta undersökningar.

#### 6.2.1 Hypoteser

HYPOTES 1: Länkad data möjliggör att information kan återanvändas på nya, väntade och oväntade, sätt.

HYPOTES 2: Länkad data ger en förbättrad användarupplevelse då tjänsterna blir kraftfullare, lättare att förstå och informationen mer lättillgänglig.

HYPOTES 3: Länkad data ger en förbättrad sökmotoroptimering.

HYPOTES 4: Länkad data gör det lättare att strukturera och hålla ordning data, vilket är en fördel vid stora datamängder.

HYPOTES 5: Informationen ökar i värde då den länkas samman med andra datakällor.

HYPOTES 6: Genom att med länkad data återanvända informationsobjekt istället för att kopiera dem minimeras risken för missförstånd vid namnkonflikter.

HYPOTES 7: Länkad data möjliggör för en semantisk webb, där maskiner förstår innebörd och relationer istället för enstaka ord.



# Referenser

- [1] Justitiedepartementet (2005-04-02). *Offentlighetsprincipen och sekretess* [www]. Hämtad 2015-06-01 från <http://www.regeringen.se/contentassets/bf7be71e6ad145e8a4e4a56ac3c3d726/offentlighetsprincipen-och-sekretess>
- [2] Regeringskansliet / Lagrummet (2010-06-03). *Lag (2010:566) om vidareutnyttjande av handlingar från den offentliga förvaltningen* [www]. Hämtad 2015-06-01 från [http://www.riksdagen.se/sv/Dokument-Lagar/Lagar/Svenskforfattningssamling/Lag-2010566-om-vidareutnyttjande-av-handlingar-fran-den-offentliga-forvaltningen/\\_sfs-2010-566/](http://www.riksdagen.se/sv/Dokument-Lagar/Lagar/Svenskforfattningssamling/Lag-2010566-om-vidareutnyttjande-av-handlingar-fran-den-offentliga-forvaltningen/_sfs-2010-566/)
- [3] P. Mankenskiöld, Sveriges kommuner och landsting. (10 December 2014). *Kommuner öppnar upp sina databaser*. [www]. Hämtad 2015-04-25 från: <http://skl.se/naringslivarbetedigitalisering/digitalisering/oppnadata/arkivoppnadata/kommuneroppnaruppsinadatabaser.4532.html>
- [4] J. Åsblom, ComputerSweden. (29 Augusti 2012). *Myndigheter går över till öppen källkod*. [www]. Hämtad 2015-04-25 från <http://computersweden.idg.se/2.2683/1.462546/myndigheter-gar-over-till-oppen-kallkod>
- [5] Open Government Working Group. (8 December 2007). *Open Government Data Principles*. [www]. Hämtad 2015-04-25 från [https://public.resource.org/8\\_principles.html](https://public.resource.org/8_principles.html)
- [6] Wikipedia. *Länkad data* [www]. Hämtad 2015-05-17 från [http://sv.wikipedia.org/wiki/L%C3%A4nkad\\_data](http://sv.wikipedia.org/wiki/L%C3%A4nkad_data)
- [7] T. Berners-Lee. (2010). *Linked Data #Is your Linked Open Data 5 Star?* [www]. Hämtad 2015-05-17 från <http://www.w3.org/DesignIssues/LinkedData.html>
- [8] opengov.se. *Öppna datakällor* [www]. Hämtad 2015-04-27 från <http://www.opengov.se/data/open/>
- [9] R. Carlsson, M. Svensson, Lunds Universitet (2013). *Länkad data som ett alternativ för att publicera öppna offentliga data* [www]. Hämtad 2015-04-27 från <http://lup.lub.lu.se/luur/download?func=downloadFile&recordId=3814405&fileId=3814410>
- [10] W3C. *Semantic web* [www]. Hämtad 2015-05-07 från <http://www.w3.org/standards/semanticweb>
- [11] W3C. *Linked data* [www]. Hämtad 2015-05-07 från <http://www.w3.org/standards/semanticweb/data>
- [12] T. Berners-Lee (2009). *The next Web of open, linked data* [www]. Hämtad 2015-05-18 från [https://www.youtube.com/watch?v=OM6XIICm\\_qo](https://www.youtube.com/watch?v=OM6XIICm_qo)

- [13] T. Berners-Lee (2006-2009). *Linked data* [www]. Hämtad 2015-05-18 från <http://www.w3.org/DesignIssues/LinkedData.html>
- [14] E. Blomqvist. *Länkade öppna data i Sverige - Portal och nationell statistik* [www]. Hämtad 2015-05-18 från <http://www.ida.liu.se/~evabl45/lodsweden.sv.shtml>
- [15] M. Palmér, H. Ebner. *Vitbok länkade öppna data* [www]. Hämtad 2015-05-19 från <http://lankadedata.se/vitbok/fordelar.html>
- [16] Länkade data i Sverige. *Länkade data* [www]. Hämtad 2015-05-19 från <http://www.linked-data.se/wiki/Info>
- [17] T. Berners-Lee (2010). *Open, Linked Data for a Global Community* [www]. Hämtad 2015-05-25 från <https://www.youtube.com/watch?v=ga1aSJXCFe0>

# Bilagor

## Bilaga 1 - Intervju med Anna Berggren, Libris

### 1: När publicerades er länkade öppna data?

*LIBRIS katalog publicerades som öppen länkad data 2008.*

### 2: Hade ni dessförinnan någon annan typ av öppna data?

*Nej*

### 3: Med vilken/vilka andra datakällor är ni sammanlänkade, och hur använder ni den informationen från dessa datakällor i er tjänst?

*LIBRIS har publicerat sitt katalogdata som öppen och länkad ända sen 2008, men det är först nu vi lagrar vårt data i grunden som länkad vilket kommer att innebära helt nya sätt att använda vpårt och andras data och berika vår katalog. Systemarkitekturen för detta är i princip färdig, men ännu inte driftsatt. Vi kommer att rulla ut delar av våra nya tjänster efter hand och i det arbetet använda oss av externa källor för att tex berika vårt katalogdata och våra söktjänster. Exempel på datakällor vi kommer att använda oss av är Dbpedia, Geodata.*

### 4: Använder dessa datakällor information från er i sina tjänster?

*Nej, vårt data är ännu inte tillgängliggjort på detta sätt.*

### 5: Har tekniken länkad data gjort att er öppna data kunnat användas på nya, oväntade vis? (Exempelvis applikationer som byggts eller nya dataset som länkat samman med er data) Om så är fallet, ge gärna exempel.

*Nej inte än, det är fortfarande för tidigt att säga, men förhoppningen är förstås att detta kommer att ske.*

### 6: Upplever ni att sammanlänkningen av er data med andra dataset gjort att värdet på er data ökat?

*Inte än*

### 7: Upplever ni att tekniken länkad data gjort eran tjänst (för användarna) -

- kraftfullare?
- lättare att förstå?
- i stort till en bättre användarupplevelse?

**Beskriv gärna hur responsen från användarna varit?**

*Går ännu inte att svara på.*

### 8: Upplever ni att teknologin länkad data förbättrat er synlighet hos sökmotorer?

*Vet inte ännu*

### 9: Har teknologin länkad data gett er några fördelar internt i organisationen?

**(Exempelvis bättre struktur på informationen)**

*Inte än. I vår organisation har vi stora datamängder som är beskrivna i flera olika format. Vi har utvecklat en systemarkitektur som är formatoberoende, men de olika formaten måste fortfarande förstå varandra för att det ska gå att använda olika datakällor och på så sätt förbättra våra tjänster, såväl externa som interna. Inom vårt team jobbar vi i första hand med nationella tjänster och vårt första mål är att översätta ett biblioteksspecifikt format till standardiserade beskrivningsformer. När detta arbete är avklarat kan vi fortsätta att mappa även andra format för att på så sätt kunna få olika datakällor att berika varandra.*

**10: Har tekniken länkad data möjliggjort att ni kan återanvända information istället för att kopiera den?**

**Om så är fallet, upplever ni att det underlättat er tjänst och/eller arbetet för er i fråga om:**

- Att undvika missförstånd (det vill säga att det blir tydligt att det "pekas på" samma sak)
- Arbets- och tidsbesparingar?

*Inte än. Målet är att vi ska undvika redundans i såväl data som i arbetsflöden. Målet är att effektivisera våra arbetsprocesser och dessutom förbättra datakvaliteten i våra tjänster.*

**11: En fördel som beskrivs med länkad data är att dess standarder möjliggör för maskiner att förstå innebörden av informationen. Har ni i er tjänst dragit några fördelar av detta?**

**Om så är fallet, beskriv gärna hur.**

*Vi mappar fortfarande vårt data från det biblioteksspecifika formatet det är beskrivet i idag. Genom att använda standardiserade beskrivningsformer kan vi tillgängliggöra biblioteksdata för internet och maskin.*

**12: Skulle det för er innebära någon skillnad om fler myndigheter publicerade sin data som länkad? (Exempelvis att ni isåfall vill bygga vidare på er tjänst ytterligare)**

*Ja, vi skulle kunna använda auktoritetsdata som andra myndigheter är specialister på och därmed förbättra datakvaliteten och effektivisera vårt arbete.*

**13. Finns något negativt att säga om tekniken länkad data?**

**(Exempelvis att det inneburit problem för er eller för era användare)**

*Nej inte än. Ska dock inte vara ngn övertro på det länkade formatet, det måste användas rätt för att få effekt.*

**14: Något övrigt ni vill tillägga om era erfarenheter kring länkad data som inte tagits upp i frågorna?**

-

## Bilaga 2, Intervju med Johan Carlström, Riksantikvarieämbetet

### 1. När publicerades er länkade öppna data?

2008-2010 drev vi ett utvecklingsprojekt (K-samsök) som var ett regeringsuppdrag och gick ut på att skapa en tjänst som skördar data från olika minnesinstitutioner (ffa museer) och RAÄ och tillgängliggjorde dem i RDF/XML-format via ett API. Datan mappades mot ett gemenast protokoll vilket skapade nya möjligheter att koppla ihop data och samsöka. De första dataseten (bl a fornminnesdata från RAÄ) publicerades 2009. 2009 var alltså året då vi publicerade länkad öppen data för första gången. Efter det har mer data tillkommit och ligger idag på ca 6 miljoner objekt totalt från ett 50-tal institutioner. Det är således en lite speciell lösning där vi publicerar både vår egen och andras data som rdf.

### 2. Hade ni dessförinnan någon annan typ av öppna data?

Kring samma tid (2009) började vi jobba med att undersöka hur vi bäst skulle kunna jobba med öppna licenser för att tydliggöra vad andra får göra med datan. Som ett test började vi publicera public domain-bilder på Flickr (<http://www.flickr.com/commons>) ungefär samtidigt. Man har även sedan tidigare kunnat göra datauttag från FMIS, dvs fornminnen, i form av shape- och tab-filer.

### 3. Med vilken/vilka andra datakällor är ni sammanlänkade, och hur använder ni den informationen från dessa datakällor i er tjänst?

Mycket av länkningen görs via dataseten inom K-samsök. Vi har även sparat ner användarskapade länkar (manuellt eller maskingenererade) länkar mellan K-samsök och litteratur från Libris, artiklar i Wikipedia och bilder i Wikimedia Commons. Datan är sökbar via ett API och används t ex i Kringla, exempel: <http://www.kringla.nu/kringla/objekt?referens=raa/fmi/10048200010001>

### 4. Använder dessa datakällor information från er i sina tjänster?

Wikipedia och Wikimedia Commons använde de persistenta länkarna i K-samsök för att länka mellan t ex artikeln om Rökstenen och beskrivningen i Fornsök, exempel: <http://sv.wikipedia.org/wiki/R%C3%B6kstenen>

### 5. Har tekniken länkad data gjort att er öppna data kunnat användas på nya, oväntade vis? (Exempelvis applikationer som byggts eller nya dataset som länkat samman med er data) Om så är fallet, ge gärna exempel.

Ffa fornminnesdata har använts i ett antal appar (se <http://www.ksamsok.se/goda-exempel/>) men inga är särskilt oväntade.

### 6. Upplever ni att sammanlänkningen av er data med andra dataset gjort att värdet på er data ökat?

Informationen sätts i ett sammanhang på ett annat sätt än tidigare vilket ökar förståelsen och nyfikna användare har möjlighet att klicka sig vidare.

### 7: Upplever ni att tekniken länkad data gjort eran tjänst (för användarna) -

- kraftfullare?
- lättare att förstå?
- i stort till en bättre användarupplevelse?

**Beskriv gärna hur responsen från användarna varit?**

*Se fråga 6.*

**8. Upplever ni att teknologin länkad data förbättrat er synlighet hos sökmotorer?**

*Svårt att säga, troligtvis inte.*

**9. Har teknologin länkad data gett er några fördelar internt i organisationen?**

**(Exempelvis bättre struktur på informationen)**

*Arbetet med länkade öppna data har ändrat vårt perspektiv och ser information i förhållande till andra datamängder, både nationellt och internationellt. Det har även gjort att vi blivit mer medvetna om standarder och masterdata.*

**10. Har tekniken länkad data möjliggjort att ni kan återanvända information istället för att kopiera den?**

**Om så är fallet, upplever ni att det underlättat er tjänst och/eller arbetet för er i fråga om:**

- Att undvika missförstånd (det vill säga att det blir tydligt att det "pekas på" samma sak)
- Arbets- och tidsbesparingar?

*Jag har inte så många konkreta exempel men vi använder t ex länkar som från bilders metadata (som beskriver fornlämningar) för att länka in bilderna i Fornsök, exempel*

<http://kulturarvsdata.se/raa/fmi/html/10094200530001>

**11. En fördel som beskrivs med länkad data är att dess standarder möjliggör för maskiner att förstå innebörden av informationen. Har ni i er tjänst dragit några fördelar av detta?**

**Om så är fallet, beskriv gärna hur.**

*Inte vad jag vet*

**12. Skulle det för er innebära någon skillnad om fler myndigheter publicerade sin data som länkad? (Exempelvis att ni isåfall vill bygga vidare på er tjänst ytterligare)**

*Det finns flera myndigheter som Riksantikvarieämbetet har kopplingar till, t ex Riksarkivet, Lantmäteriet och Länsstyrelsen, vars data skulle vara både intressant och nödvändig att kunna länka till och återanvända.*

**13. Finns något negativt att säga om tekniken länkad data? (Exempelvis att det inneburit problem för er eller för era användare)**

*Tekniken är förhållandevis ny och används inte i så stor utsträckning inom offentlig sektor. Därmed svårt att jämföra och se det långsiktiga i att förvalta denna typ av tekniker.*

**14. Något övrigt ni vill tillägga om era erfarenheter kring länkad data som inte tagits upp i frågorna?**

*Nej*