

# **METADATA ENLIGT DUBLIN CORE**

**Tillämpningar och konsekvenser i de svenska  
kvalitetssöktjänsterna SAFARI, Svenska miljönätet  
och Svesök.**

**Marita Fagerlind  
Gunilla Gisselqvist**

Examensarbete (20 poäng) för magisterexamen i Biblioteks- och informationsvetenskap vid Lunds universitet.

Handledare: Birgitta Olander, Jörgen Eriksson, Colm Doyle

BIVILs skriftserie 1999:11

ISSN 1401-2375

© Lunds universitet. Biblioteks- och informationsvetenskap 1999

## **ABSTRACT**

This study investigates the individual usage of the Dublin Core metadata format in three Swedish Subject-Based-Information-Gateways, SAFARI, Svenska miljönätet and Svesök.

The authors also present the reader with an overview of the metadata concept and explains in particular the Dublin Core Element Set.

The approach to the study is investigation of Dublin Core application from the viewpoint of three major issues, user friendliness, interoperability and quality control. The study includes an in-depth comparison of the three different metadata templates, as well as interviews with people in charge of the metadata production at the respective SBIGs and users of the respective SBIGs' templates.

The study shows that the three SBIGs are more similar than they are different. They have all succeeded in designing their templates in a user-friendly way. The relatively minor differences between the three will most likely not interfere with their mutual interoperability.

In conclusion, the authors find that a quality controlled web environment appears imperative to ensure the success of the Dublin Core metadata format. In an environment of this kind, the Dublin Core format, with its general appeal of simple yet adequate resource description, may very well change the structure of Internet resources for the better.

## FÖRORD

Vi vill inleda med att tacka de personer som hjälpt oss i arbetet med uppsatsen. Först och främst vår handledare Birgitta Olander för det stöd och den inspiration hon gett oss. Ett stort tack också till Jörgen Eriksson och Colm Doyle på LUB NetLab, utan vars kunskaper vi hade haft svårt att sätta oss in i ämnet DC-metadata.

Vi vill också tacka ansvariga för de söktjänster vi samarbetat med under arbetets gång. Martin Sparr på SAFARI, Annica Moshiri på Svenska miljönätet och Lotta Jansson, Anette Svensson och Christina Törnqvist-Westin på Svesök. Det har varit en glädje att arbeta tillsammans med personer som varit så tillmötesgående och positivt inställda till vårt arbete.

Slutligen tackar vi de metadataproducenter och alla andra inblandade som ställt upp och svarat på våra e-postfrågor.

Lund 1999-06-29

Marita Fagerlind och Gunilla Gisselqvist

# INNEHÅLLSFÖRTECKNING

<b>1. INLEDNING</b> .....	<b>5</b>
<b>2. SYFTE OCH FRÅGESTÄLLNINGAR</b> .....	<b>7</b>
<b>3. METOD OCH AVGRÄNSNINGAR</b> .....	<b>9</b>
<b>4. LITTERATURGENOMGÅNG</b> .....	<b>12</b>
4.1 METADATA.....	12
4.2 THE DUBLIN CORE ELEMENT SET .....	15
<b>4.2.1 Qualifiers</b> .....	19
<b>4.2.2 RDF och The Warwick Framework</b> .....	21
<b>4.2.3 Dublin Core Workshops</b> .....	21
4.3 SAMORDNINGSFÖRSÖK .....	25
4.4 Z39.50 .....	25
4.5 DUBLIN CORE I FRAMTIDEN.....	27
<b>5. PRESENTATION AV SÖKTJÄNSTERNA</b> .....	<b>28</b>
5.1 SAFARI.....	28
<b>5.1.1 Innehåll och målsättning</b> .....	28
<b>5.1.2 Kvalitetskriterier</b> .....	29
<b>5.1.3 Inmatningsformulär</b> .....	29
5.2 SVENSKA MILJÖNÄTET .....	31
<b>5.2.1 Innehåll och målsättning</b> .....	31
<b>5.2.2 Kvalitetskriterier</b> .....	32
<b>5.2.3 Inmatningsformulär</b> .....	33
5.3 SVESÖK.....	34
<b>5.3.1 Innehåll och målsättning</b> .....	34
<b>5.3.2 Kvalitetskriterier</b> .....	35
<b>5.3.3 Inmatningsformulär</b> .....	36
<b>6. RESULTATREDOVISNING</b> .....	<b>38</b>
6.1 SAMORDNINGSFÖRSÖK .....	38
<b>6.1.1 Val av metadataformat</b> .....	38
<b>6.1.2 Lokala varianter och språkbruk</b> .....	39
<b>6.1.3 Inställning till interoperabilitet</b> .....	52
<b>6.1.4 Sökmöjligheter</b> .....	52
<b>6.1.5 Granularitet</b> .....	54
6.2 ANVÄNDARVÄNLIGHET.....	54
<b>6.2.1 Formulärens utformning</b> .....	54
<b>6.2.2 Dialog med användarna</b> .....	56
<b>6.2.3 Hjälpfunktioner och användarutbildning</b> .....	56
<b>6.2.4 Användarnas upplevelser</b> .....	56
<b>6.2.5 Webbproducentens insatser vad beträffar inmatning av metadata</b> .....	58
6.3 KVALITETSKONTROLL .....	59
<b>6.3.1 Kontroll av innehåll och metadata</b> .....	59
<b>7. ANALYS OCH DISKUSSION</b> .....	<b>60</b>
7.1 SAMORDNINGSFÖRSÖK .....	60
7.2 ANVÄNDARVÄNLIGHET.....	64
7.3 KVALITETSKONTROLL .....	67
<b>8. SLUTSATSER</b> .....	<b>69</b>

**ORDLISTA OCH FÖRKORTNINGAR.....71**

**KÄLLFÖRTECKNING.....74**

**BILAGOR**

- Bilaga 1: SAFARI:s inmatningsformulär
- Bilaga 2: Svenska miljönätets inmatningsformulär
- Bilaga 3: Svesöks inmatningsformulär
- Bilaga 4: Metadata om uppsatsen

## 1. INLEDNING

Orsaken till valet av uppsatsämne är vårt relativt nyväckta intresse för elektronisk informationsutveckling. Vi är båda flitiga Internetanvändare och fascineras av de möjligheter detta medium erbjuder. Samtidigt har vi, liksom många andra, reagerat över den brist på struktur som råder på nätet och som i sin tur leder till irritation, tidsspillan och ständiga missförstånd. Detta märks inte minst när man försöker hitta information genom de många söktjänster som finns på nätet. Med sökningarna tillkommer en ansemlig mängd brus samt onödiga upprepningar av olika sidor inom samma sajt i träfflistorna. Vi tyckte Martin Sparr i söktjänsten SAFARI uttryckte oordningen på Internet på ett målande sätt:

“Ordning på Internet, behövs det? Tänk om ett bibliotek tillät alla världens författare att ställa in sina egna böcker på vilken hylla som helst och dessutom ändra i innehållet lite då och då och ställa böckerna på olika hyllor från ett år till ett annat. Vilken röra det skulle bli! Det är därför lite konstigt att vi tillåter detta på Internet.”<sup>1</sup>

Blivande bibliotekarier och informationsvetare som vi är, bestämde vi oss för att undersöka något som skulle kunna bringa ordning i detta kaos. Det skulle alltså vara önskvärt med någon sorts ordning på Internet motsvarande den som finns på bibliotek. På bibliotek katalogiseras varje bok och annat medium av en bibliotekarie. Det är dock omöjligt att göra samma sak på Internet. Inget bibliotek eller annan institution har tillräckliga resurser för att ta itu med denna oerhört omfattande marknad av information. I våra efterforskningar kom vi dock i kontakt med uttrycket metadata och insåg att detta fenomen hade stora förväntningar på sig att uppfylla just det vi, och många med oss, funderat över. Vi bestämde oss för att fördjupa oss i ämnet och valde det format som kallas Dublin Core, eftersom det är så nära en standard man kan komma idag.

Metadata går i korta drag ut på att katalogisera elektroniska dokument på Internet och idén med Dublin Core handlar om att alla som skapar ett elektroniskt dokument också ska kunna märka det med metadata och på så vis ordna själv med katalogiseringen. Detta tjänar alla parter på. Informationssökaren hittar de dokument som intresserar honom/henne och dokumentproducenten når fram med den information han/hon vill förmedla. Bruset kommer att avta avsevärt.

Det tog oss en ansemlig tid att sätta oss in i ämnet, som vi till en början fann svårgenomträngligt. Vi har stött på en stor mängd ord och uttryck vi inte tidigare var bekanta med och som vi i vår tur haft vissa besvär med att återge i begripliga termer för icke insatta i ämnet. Det är vår förhoppning att både specialintresserade och mindre bevandrade inom området ska kunna ta till sig innehållet i uppsatsen. Vi har i möjligaste mån försökt förmedla ämnet på ett enkelt och lättfattligt sätt. Det finns enligt våra erfarenheter en djup klyfta mellan professionella inom ämnet och lekmän. Eftersom syftet med DC-formatet är

---

<sup>1</sup> Sparr, Martin. SAFARI. <URL: <http://www.hsv.se/verksamhet/utredningar/safari/metadata.html>> 981130

att gemene man ska kunna lära sig att märka sina dokument med metadata, tycker vi det är av värde att förmedla kunskap om ämnet och så långt det är möjligt avmystifiera det.

Det finns fler metadataformat än Dublin Core. Vi kommer dock inte beskriva dessa p.g.a. tidsbrist. De intresserade kan läsa om dem i Nordiska metadataprojektet<sup>2</sup>.

Vi vill förtydliga att vi i uppsatsen synonymt använder oss av uttrycken Internet, nätet, www och webben. Vi är dock medvetna om skillnaden i betydelse mellan begreppen. Den har vi redogjort för i ordlistan.

Dublin Core är ett metadataformat anpassat för internationella förhållanden, varför s.k. elementnamn, qualifiers m.m. är på engelska. Det finns i dagsläget ingen vedertagen svensk översättning av DC-termerna. De projekt vi undersökt i uppsatsen har i vissa fall översatt termerna till svenska. Detta har resulterat i flera varianter av översättningar. Vi har därför valt att så konsekvent som möjligt använda de vedertagna engelska uttrycken i våra diskussioner.

Nämnvärt är att Svenska miljönetets inmatningsformulär<sup>3</sup> ändrades i slutskedet av uppsatsskrivandet. Utskrifter av de tre söktjänsternas inmatningsformulär, såsom de såg ut HT 98, återfinns därför som bilagor.

---

<sup>2</sup> Hakala, Juha. Medförfattare: Hansen, Preben. Husby, Ole. Koch, Traugott och Thorborg, Susanne. *The Nordic metadata project Final report*. 1998. Kapitel 1.2.1 "Evaluation of existing metadata formats."

<URL: <http://linnea.helsinki.fi/meta/nmfinal.htm>> 981130

<sup>3</sup> Se ordlistan.

## 2. SYFTE OCH FRÅGESTÄLLNINGAR

Metadata röner idag stort intresse, både i Sverige och utomlands. Formatet har dock ännu inte nått någon avsevärd utbredning, utan tillämpas idag mestadels i mindre projekt. Tanken och förhoppningen inom branschen är dock att metadata ska fungera i ett större sammanhang på Internet så småningom. För att detta ska vara ekonomiskt och praktiskt genomförbart krävs att webbproducenten själv märker sina dokument med metadata. Detta innebär en stor frihet. Webbproducenten kan i princip använda vilka ämnesord som helst för att beskriva sin resurs. Missbruk av metadata resulterar i att träffarna fortfarande inte blir logiska och brusfaktorn i princip riskerar att bli densamma som utan metadata. Risk för missförstånd föreligger också. När man första gången använder sig av ett inmatningsformulär är det inte självklart hur man korrekt ska fylla i det. För att undvika missbruk och missförstånd krävs någon form av kvalitetskontroll.

Med denna diskussion som bakgrund har vi i uppsatsen studerat tre svenska kvalitetssöktjänster. Syftet med undersökningen har varit att försöka ta reda på hur söktjänsterna förhåller sig till dessa tre aspekter:

### **Samordningsförsök** (interoperabilitet<sup>4</sup>)

DC strävar efter att bli en standard och själva meningen med en standard är att den ska kunna fungera i ett större sammanhang, alltså att den ska kunna appliceras i flera söktjänster samtidigt. Skulle det vara möjligt att samsöka i de tre söktjänsterna idag och vad görs för att detta ska bli en realitet?

### **Användarvänlighet** (användare = de som matar in metadata)

Om metadata ska fungera i ett större sammanhang på nätet krävs att webbproducenten själv märker sina dokument med metadata. Vilka är fördelarna respektive nackdelarna med detta? Vad har söktjänsterna gjort för att underlätta för användarna?

### **Kvalitetskontroll** (både vad beträffar innehåll och metadatamärkning)

Syftet med metadata är att informationssökaren ska få relevanta träffar på sina sökfrågor, alltså är kvalitetskontroll av både innehåll och metadatamärkning av yttersta vikt. Det gäller att undvika missbruk och missförstånd. Har söktjänsterna möjligheter att lägga ner resurser på kontroll av innehåll och metadatamärkning och i så fall i vilken utsträckning?

Frågeställningarna lyder som följer:

- Hur har de undersökta söktjänsterna valt att utforma sina inmatningsformulär? Vad har detta fått för konsekvenser vad beträffar samsökningar?
- Hur användarvänligt är inmatningsformulären utformade?
- Existerar någon form av kvalitetskontroll inom söktjänsterna?

---

<sup>4</sup> Se ordlistan.



- Hur skulle DC-metadatan kunna fungera i ett större sammanhang på Internet?

Vårt syfte har också varit att lättfattligt men likväl relativt ingående förklara begreppen metadata och Dublin Core. Detta i ett försök att ge den oinvidigade en bakgrund i ämnet samt att ge en inblick i den aktuella diskussionen.

### 3. METOD OCH AVGRÄNSNINGAR

Begreppet metadata var oss inte helt obekant när vi valde uppsatsämne. Våra kunskaper var dock högst begränsade, varför vi fick lägga ner mycket tid på att sätta oss in i ämnet och på att hitta lämpliga resurser, som dessutom innefattade ett språk som vi kunde ta till oss. Det mesta vi läst har varit på engelska och skrivet på ett språk som innefattat fackuttryck som vi haft svårt att översätta till motsvarande svenska termer. Många resurser om metadata är dessutom skrivna på ett tekniskt komplicerat sätt och eftersom vi är humanister hade vi vissa svårigheter att sätta oss in i denna nya begrepps värld. Det som intresserade oss var metadata som informationsstruktur, men ofelbart strävade den utvalda litteraturen mot beskrivning av tekniken bakom, där vi till en början kände oss bortkomna. Allt eftersom uppsatsen tog form har vi dock lyckats tillskansa oss vissa tekniska kunskaper. Vi har sedermera ansträngt oss att återge dessa på en nivå som även den tekniskt oinvidde ska kunna ta till sig. Vi har dessvärre sett oss nödgade att begränsa den tekniska beskrivningen något. De vanligast förekommande företeelserna har vi satt oss in i och lagt vikt vid att beskriva, men försökt att inte gå för djupt in på detaljer, som i sin tur ofelbart skulle lett till ytterligare krav på teknisk beskrivning.

Ämnet metadata är under utveckling, varför vi huvudsakligen koncentrerade oss på att hitta litteratur på Internet. Tryckt litteratur innebär ofta att ämnet redan är på väg in i nästa utvecklingsfas, då man efter tryckning och utgivning slutligen läser verket. Vi startade alltså med att grundligt läsa om metadata i allmänhet och DC-metadata i synnerhet med hjälp av Internetresurser. Det har varit intressant att läsa om ämnet på Internet av den anledningen att informationen har varit ny eller nyligen uppdaterad. Vi har verkligen känt att vi befunnit oss mitt i en utvecklingsprocess och märkt hur innehållet ändrat karaktär, ibland bara över en natt. Detta har samtidigt vållat oss visst huvudbry, då vi känt oss manade att hela tiden ha den snabba utvecklingen under uppsikt. Ett annat problem med att övervägande läsa Internetresurser var att läsoplevelsen ofta blev splittrad, då nätet hela tiden erbjuder referenser hit och dit, vilka vi kände oss manade att följa upp och undersöka. Det hände att vi gick vilse bland länkarna och inte hittade tillbaka till utgångsresursen. Vi anser detta till trots att Internetresurserna varit ovärderliga och valt att fokusera mest på dem. Vi har även till viss del läst böcker och tidskriftsartiklar. Problemet med snabb utveckling kontra dålig uppdatering gjorde sig dock allt som oftast påmint.

I litteraturgenomgången har vi inte refererat till alla källor vi läst inom området, då mycket av informationen återkommit gång på gång. Vi har valt att basera innehållet på några få men omfattande och informativa källor.

Orsaken till att vi koncentrerat oss på formatet Dublin Core är att vi genom våra studier uppfattat att detta är så nära en standard man kan komma idag. DC är det mest utbredda av metadataformaten och i dagsläget har ganska många projekt runt om i världen börjat tillämpa det för katalogisering av elektroniska resurser. Eftersom vi samarbetade med NetLab<sup>5</sup>, som är engagerade i DC:s utveckling, föll också valet naturligt på detta format. DC är också särskilt utbrett i Norden, vilket var ytterligare en bidragande orsak till vårt val.

---

<sup>5</sup> Se ordlistan.

Vi hade klart för oss att vi ville undersöka hur DC fungerar i praktiken och bestämde oss sålunda för att närmare studera ett antal projekt som valt att tillämpa formatet. Genom resurslistor på Internet lyckades vi hitta ett trettiotal internationella metadata-databaser, som vi systematiskt gick igenom för att hitta lämpliga studieobjekt. Anledningen till att vi valde bort flertalet av dessa beror på flera faktorer. Projekten var uppbyggda på vitt skilda sätt. Vissa använde sig av något annat metadataformat än DC. Flera projekt var uppbyggda på ett sätt som vi upplevde som svårtillgängligt och komplicerat. Vidare innehöll vissa varken någon söktjänst eller länksamling, varför vi hade svårt att förstå syftet med projekten. Somliga var dessutom så pass nystartade och befann sig på ett utvecklingsstadium, som gav ett ostrukturerat och ibland kaotiskt intryck och i förlängningen skulle göra en undersökning komplicerad. Vidare fanns det de projekt som använde sig av DC:s grundidé utan någon egen variant alls, varför hela poängen med vår undersökning förfelades, då syftet med den var att göra en jämförande studie mellan olika lokala varianter av DC-tillämpning. I vissa fall lyckades vi inte ens ta reda på vilket inmatningsformulär som tillämpades. Några projekt använde sig visserligen av DC, men kombinerade detta med ett eller flera andra metadataformat, vilket gav ett mycket omfattande material att bearbeta och som av den anledningen avskräckte oss.

Efter denna första utrensning kvarstod ett antal svenska, europeiska och australiensiska projekt. Dessa gick vi igenom relativt detaljerat. De utländska projekten var:

- Ariadne<sup>6</sup>
- EdNA<sup>7</sup>
- Mathnet<sup>8</sup>
- NewsAgent for Libraries<sup>9</sup>
- SSG-FI<sup>10</sup>
- Medical Metadata Projekt<sup>11</sup>

Dessa projekt var samtliga intressanta ur metadata-synpunkt. De var dock utformade så pass olika samt att vissa innehöll DC-metadata som skilde sig en hel del från den standard vi studerat. Ur detta perspektiv fann vi det svårt att närmare analysera projekten. Vi skickade dock e-post till ansvariga för vissa av projekten med intresseförfrågningar, men fick ingen respons, varför vi valde bort även dessa. Vi koncentrerade oss istället på Sverige och detaljstuderade ett antal projektbeskrivningar på Internet och blev glatt överraskade över deras välstrukturerade samlingar samt relevanta tillämpning av DC-elementen. Vi bestämde oss slutligen för att koncentrera oss på dessa tre svenska kvalitetssöktjänster:

- SAFARI<sup>12</sup>. Spridning av forskningsinformation till allmänheten över Internet. Svenska resurser.

---

<sup>6</sup> <URL: <http://ariadne.unil.ch/>> 981030

<sup>7</sup> <URL: <http://www.otfe.vic.gov.au/edna/dc5edna.htm>> 981030

<sup>8</sup> <URL: <http://www.math-net.de/>> 981030

<sup>9</sup> <URL: <http://www.sbu.ac.uk/litc/newsagent/>> 981030

<sup>10</sup> <URL: <http://www.sub.uni-goettingen.de/ssgfi/>> 981030

<sup>11</sup> <URL: <http://medir.ohsu.edu/~maletg/MedMetadata.HTM>> 981030 (död länk 990121)

<sup>12</sup> <URL: <http://www.hsv.se/verksamhet/utredningar/safari/index.html>> 981230

- Svenska miljönätet<sup>13</sup>. Miljödata, miljöstatistik och miljöinformation med relevans för miljöarbetet i Sverige.
- Svesök<sup>14</sup>. KB:s rikstäckande söktjänst som främst vänder sig till biblioteken.

Vi kontaktade ansvariga för respektive projekt med förfrågan om intresse för deltagande i vår undersökning samt intervjuer fanns, och fick omedelbar och positiv respons från samtliga.

På ett tidigt stadium hyste vi förhoppningar om att arbeta med och jämföra ett tiotal projekt. Allt eftersom undersökningen av de potentiella studieobjekten fortlöpte, insåg vi det orealistiska i att ta på oss ett allt för stort antal. Resultatet riskerade att bli splittrat och endast en ytlig undersökning av varje projekt skulle kunna komma till stånd. Valet av de tre svenska välstrukturerade och genomtänkta söktjänsterna gav oss en möjlighet att i detalj förkovra oss i respektive projekt.

Undersökningen fortlöpte sedan på följande sätt; dels detaljstuderade vi projektens inbördes söktjänster, länksamlingar, metadataformulär samt läste deras projektbeskrivningar på Internet, dels utformade vi e-postintervjuer till respektive söktjänst. Dessa formulerades delvis olika, då projektbeskrivningarna på Internet gav oss olika mycket information, vilket resulterade i att vi fick fråga om de luckor i informationen som uppstod vid studierna av projektbeskrivningarna. Intervjuerna är alltså ett led i vår faktainsamling. E-postintervjuerna fick vi omedelbar respons på och har funnit samtliga tre projekt mycket givande att arbeta med. De tre utvalda projekten lämpade sig väl för en undersökning, då de var relativt stora och etablerade. Visserligen ganska nystartade, men ändå med tillräckligt mycket resurser för att fungera i ett analyserande perspektiv. Alla tre hade dessutom utformat sina egna varianter av metadataformulär, vilket gjorde en jämförelse dem emellan intressant. Vi fördjupade oss i hur användarvänligt inmatningsformulären hade utformats i respektive söktjänst, genom att bl.a. genomföra en användarundersökning. Vi nöjde oss med att skicka ut en kortare e-postenkät till några av dem som har använt sig av söktjänsternas inmatningsformulär. Av svaren valde vi att bearbeta fyra enkäter inom varje projekt. Det hade varit värdefullt med en mer omfattande undersökning vad beträffar användarvänlighet, eftersom det är av största vikt att inmatningsformulären anpassas för användarnas ändamål, men p.g.a. tidsbrist och ett redan ganska omfattande material, var vi tvungna att begränsa undersökningen. Önskvärt hade också varit att göra en omfattande undersökning om hur väl träffarna stämmer överens med sökfrågorna. Av ovan nämnda orsak har vi även här valt att göra en avgränsning.

---

<sup>13</sup> <URL: <http://smn.viron.se/>> 981230

<sup>14</sup> <URL: <http://www.svesok.kb.se/>> 981230

## 4. LITTERATURGENOMGÅNG

### 4.1 METADATA

Som alla som har provat vet, är det svårt att hitta relevanta dokument i en sökning på nätet. Webben är stor och det finns ingen ”hylluppställning”, inga kataloger och ingen uppenbar ordning. Paul Miller skriver i artikeln ”Metadata for the masses” om hur sökningarna på Internet blir fel då man inte kan kategorisera sökorden, d.v.s. kunna bestämma om ordet man söker på är t.ex. ett personnamn, en ort eller en tidskrift<sup>15</sup>. Om man t.ex. vill hitta en avhandling, och allt man vet är att författaren heter Lund i efternamn, och söker på ”Lund” i AltaVista får man 380863 träffar som handlar om allt från tiderna för Flygbussarna Malmö-Lund till vad de tror på i Pingstkyrkan i Lund. Sökmotorerna lyckas producera massor av resultat på en sökfråga, men det är däremot svårt att få relevanta träffar. Och för att få just relevanta träffar måste man kunna kategoribestämma sökningarna d.v.s. ”tala om” för söktjänsten att det är just ett personnamn man letar efter. För att kunna göra denna urskiljning i sökningarna måste det finnas motsvarande urskiljningar i beskrivningarna av dokumenten som sökmotorn letar bland, annars blir det ingen träff. Hur gör man då en sådan dokumentbeskrivning? Jo, med hjälp av metadata.

Ordet ”meta” betyder diskussion kring den egna formen. Metadata betyder således data om data. Man kan kalla det en form av katalogisering. Metadata har dock kommit att handla om data om elektronisk data, d.v.s. katalogisering av webbresurser. Det finns ingen standardiserad definition av ordet. Roy Tennant beskriver det i en artikel som strukturerad information om information<sup>16</sup>, medlemmar i UKOLN Metadata Group kallar det data som beskriver resursens egenskaper eller utmärkande drag<sup>17</sup>, Tim Berners-Lee definierar det som maskinbegriplig information om webbresurser och annat<sup>18</sup> och Priscilla Caplan anser att det helt enkelt är ett annat ord för katalogisering<sup>19</sup>. En sak kan man dock konstatera, metadata och katalogisering är inte riktigt synonyma begrepp. Katalogisering är ett ord som hör hemma speciellt i bibliotekskretsar och innefattar tryckt eller åtminstone fysiskt material, medan metadata är resursbeskrivning av elektroniska objekt, företrädesvis på Internet. Metadata har däremot ingenting med metasökverktyg eller metasöktjänst att göra, då dessa är sökverktyg eller söktjänster vars sökrobotar söker i flera databaser än en. Meningen med metadata, hur man än definierar ordet, är informationsbeskrivning och informationsåtervinning, d.v.s. att beskriva ett objekt på ett sådant sätt att det så lätt som möjligt går att återfinna, precis som med traditionell katalogisering. Ett resultat av metadata, och den huvudsakliga anledningen till att det finns metadata, är alltså att förbättra sökbarheten för innehållet på Internet. Man kan säga att metadata är ett sätt att göra data på webben till information. Dock kan man inte enbart hänvisa till den ordning som

<sup>15</sup> Miller, Paul. *Metadata for the masses*. 1996. <URL: <http://www.ariadne.ac.uk/issue5/metadata-masses/>> 981030

<sup>16</sup> Tennant, Roy. ”21st Century Cataloging”. *Library Journal*. 123(7) 15 april 1998, s.30.

<sup>17</sup> UKOLN (UK Office of Library Networking) Metadata Group. *An Overview of Resource Description Issues*. 1998. <URL: [http://www.ukoln.ac.uk/metadata/desire/overview/rev\\_01.htm](http://www.ukoln.ac.uk/metadata/desire/overview/rev_01.htm)> 981030

<sup>18</sup> Berners-Lee, Tim. *Metadata Architecture*. 1998. <URL: <http://www.w3.org/Designissues/Metadata.html>>981030

<sup>19</sup> Caplan, Priscilla. ”You call it corn, we call it syntax-independent metadata for document-like objects”. *The Public-Access Computer Systems Review* 6. 1995(4). <URL: <http://info.lib.uh.edu/pr/v6/n4/capl6n4.html>> 981030

råder på ett bibliotek när man refererar till den ordning som råder på Internet. Internet är ju faktiskt inte ett enda stort virtuellt bibliotek, vilket är ganska lätt att uppfatta det som. En av de stora skillnaderna är att Internet inte är centralstyrt. Så som bibliotekarier sedan länge katalogiserat nyinkomna böcker, så kan ingen katalogisera nya dokument på Internet. Dock behövs bibliotekariers kunskap i sammanhanget. Clifford Lynch skriver i sin artikel "Searching the Internet" att Internet är en kaotisk blandning av allt möjligt som går att lägga ut på nätet, förutom textdokument, bilder, rapporter, menyer, reklam o.s.v. Fortsätter det att växa och frodas som ett nytt kommunikationsmedel, påpekar han vidare, så kommer det traditionella bibliotekets kunskaper om hur man organiserar information behövas. Men även om så är fallet kommer nätet dock inte att likna ett traditionellt bibliotek eftersom innehållet är mer vidgående än en standardiserad bibliotekssamling. Därför, menar Lynch, behövs både bibliotekarier och tekniker i samarbete kring utvecklingen av Internet<sup>20</sup>.

"... Consequently, the librarian's classification and selection skills must be complemented by the computer scientist's ability to automate the task of indexing and storing information. Only a synthesis of the different perspectives brought by both professions will allow this new medium to remain viable."<sup>21</sup>

I dagsläget saknas en fastställd standard för metadataformat för webbresurser. Och hur skulle en sådan standard se ut? Hur bör metadata produceras och vad är viktigt att ha med i en webbresursbeskrivning? Hur specifik och detaljrik kontrasterat enkel och lättfattlig bör den vara? I biblioteksvärlden finns MARC-katalogisering vilken är sig ganska lik världen över men något sådant finns inte på Internet, förutom helt generellt att metadatabeskrivningar ska klistras in under namnet "META NAME" mellan <HEAD> och </HEAD> i ett dokument's HTML-kod<sup>22</sup>.

Paul Miller skriver att många av de metadataformat som finns idag är så specialiserade och komplicerade att de inte passar som beskrivning av allmänna resurser<sup>23</sup>. Dublin Core är däremot en modell av metadata som kan beskriva just majoriteten av resurserna på Internet. Därför har också Dublin Core vunnit gehör internationellt och är på god väg att bli en standard. Martin Sparr på SAFARI spår att Dublin Core "...blir sannolikt standard inom en snar framtid, vilket kommer att öka användningen ytterligare."<sup>24</sup> En anledning kan vara att Dublin Core har uppstått just kring samarbete mellan bl.a. bibliotekarier och tekniker. Men kanske framför allt för att det är användbart för att beskriva så många olika typer av objekt. Alla s.k. "dokument-liknande objekt"<sup>25</sup> kan beskrivas med Dublin Core-metadata, d.v.s. inte endast textdokument, utan också bilder, musikfiler m.m.

Speciellt bibliotek och museer har visat stort intresse för Dublin Core. Metadata är alltså inte enbart en biblioteksfråga. Att t.ex. objekten på ett museum är kopplade till en textbeskrivning är ganska angeläget eftersom alla söktjänster är textbaserade.

<sup>20</sup> Lynch, Clifford. "Searching the Internet, Combining the skills of the librarian and the computer scientist may help organize the anarchy of the Internet." *Scientific American*. 1997.

<URL: <http://www.sciam.com/0397issue/0397lynch.html>> 981030

<sup>21</sup> Ibid.

<sup>22</sup> Se ordlistan.

<sup>23</sup> <URL: <http://www.ariadne.ac.uk/issue5/metadata-masses/>> 981030

<sup>24</sup> <URL: <http://www.hsv.se/verksamhet/utredningar/safari/metadata.html>> 981230

<sup>25</sup> Från engelskans "document-like objects".

UKOLN Metadata Group upp en annan och tämligen ny aspekt av metadata, en möjlighet att ha metadata om människor, om kurser, om olika forskningsinstitutioner och andra ”objekt”<sup>26</sup>. Detta för att söktjänster ska kunna leta efter resurser som matchar en speciell användarprofil, eller söka efter personer med ett speciellt forskningsintresse etc.

Ett problem med metadata är hur den ska produceras och av vem. Det finns c:a 300 miljoner ”adresserbara objekt”<sup>27</sup> på Internet idag, och antalet växer hela tiden vilket gör det ganska orimligt att någon organisation eller bibliotek skulle ta sig an det omfattande arbete det innebär att beskriva alla dokument. Det förefaller därför rimligt att anta att den som producerar ett webbdokument bör vara den som beskriver det, d.v.s. lägger in metadata om det. Det ligger trots allt i dessa s.k. dokumentproducenters intressen att deras resurser blir lättare återsökta och därmed mer lättåtkomliga. Dock riskerar man eventuellt en viss kvalitetssänkning av metadata när man låter någon som inte är insatt i begreppet katalogisering själva katalogisera. Juha Hakala, Ole Husby och Traugott Koch skriver i en artikel att även om dokumentproducenterna i stor utsträckning lägger in metadata på sina dokument, p.g.a. det stora antalet dokument, kommer specialistkatalogisatörer troligtvis att fortsätta katalogisera kvalitetsdokument med lång livslängd. Men, påpekar de, vem som än lägger in metadata är det viktigt att den överhuvudtaget skapas och att den är interoperabel<sup>28</sup>:

”...a richer metadata provision is the most important step needed to improve the quality and precision in automatic resource discovery and retrieval tools.”<sup>29</sup>

Med automatisk indexering menas alltså att det är en sökmotor som väljer ut resultaten av en sökning. I motsats till detta kan man ställa mänsklig indexering, d.v.s. när en katalogisatör gör motsvarande urskiljning ur ett material. Eftersom materialet att sortera ur, d.v.s. dokumenten på Internet, är så omfattande är det automatisk indexering som råder just nu, eftersom mänskliga insatser är kostbara i jämförelse. Det är alltså teknologin som styr hur informationen är organiserad, skriver Lynch. Men en sökmotor saknar vår mänskliga urskiljningsförmåga, varför tekniken aldrig kan garantera kvalitet av själva innehållet i dokumenten. Metadata är alltså ett sätt att försöka förbättra kvaliteten på automatgenererad indexering. Metadata kan dock aldrig ersätta mänskliga utvärderingar då dessa handlar om att göra kvalitetsurval. I hur stor utsträckning mänskliga indexeringsförmågor eller automatisk indexering behövs i framtiden kommer att bero på vem som använder Internet, fortsätter Lynch. För akademiker och forskare är modellen av en organiserad länksamling, d.v.s. ett digitalt bibliotek, relevant. För andra grupper är ett oorganiserat medium det bästa sättet att sprida och återvinna information; människor som av en eller annan anledning inte vill ha data som är ”filtrerad”<sup>30</sup>.

<sup>26</sup> <URL: [http://www.ukoln.ac.uk/metadata/desire/overview/rev\\_01.htm](http://www.ukoln.ac.uk/metadata/desire/overview/rev_01.htm)> 981030

<sup>27</sup> Från engelskans ”addressable objects”, se även ordlistan. Weibel, Stuart. *Using Web Metadata: The Dublin Core*. 1998. <URL: <http://purl.oclc.org/~emiller/talks/www7/tutorial/part1/index.htm>> 981030

<sup>28</sup> Hakala, Juha, Husby, Ole och Koch, Traugott. *Warwick Framework and Dublin core set provide a comprehensive infrastructure for network resource description Report from the Metadata Workshop II, Warwick, UK, April 1-3, 1996*. <URL: <http://www.wub2.lu.se/tk/warwick.html>> 981030

<sup>29</sup> Ibid.

<sup>30</sup> <URL: <http://www.sciam.com/0397issue/0397lynch.html>> 981030

## 4.2 THE DUBLIN CORE ELEMENT SET

Dublin Core är ett metadataformat som har utvecklats med tyngdpunkterna på att dokumentproducenten är den som själv ska lägga in metadata på de egna resurserna samt att utformningen är gjord att passa resursbeskrivningar för majoriteten av resurserna på Internet. Dublin Core vänder sig till den stora massan av dokumentproducenter och är tänkt att vara lättanvänt och lättillgängligt. Den inbyggda flexibiliteten är också viktig eftersom olika behov från användarnas sida av antingen enkla eller utförliga resursbeskrivningar ska kunna tillmötesgåas inom de gemensamma ramarna.

Dublin Core uppstod 1995 på en workshop i Dublin, Ohio USA, genom samarbete mellan bl.a. bibliotekarier, nätverkstekniker och databasspecialister. Systemet är alltså fortfarande ganska nytt och under utveckling. I Nordiska metadataprojektet berättas det dock att mellan åren 1996 och 1998 har Dublin Core stabiliserats och blivit välkänt<sup>31</sup>. Stuart Weibel på OCLC<sup>32</sup> är den som ses som DC:s drivande kraft och grundare. 1998 blev OCLC officiellt styrelse<sup>33</sup> för Dublin Core vilket betyder att OCLC officiellt står för underhållet av Dublin Core. Dublin Core har uppkommit ur ett behov av en allmän standard för elektronisk resursbeskrivning och som en kompromiss mellan alltför enkla automatiska resursbeskrivningar och manuella tidskrävande mänskliga katalogiseringsinsatser. Hakala, Husby och Koch säger vad beträffar bakgrunden till varför idén om Dublin Core uppstod att:

”Automatically generated records often contain too little information to be useful, while manually generated records are too costly to create and maintain for the large number of electronic documents currently available on the Internet. Records created from the Dublin Core are intended to mediate these extremes, affording a simple structured record that may be enhanced or mapped to more complex records as called for, either by direct extension or by link to a more elaborate record.”<sup>34</sup>

Dublin Core är tänkt att vara ett generellt och allmänt metadataformat i motsats till de komplicerade och specialiserade format som redan existerar.

OCLC har skapat en hemsida för Dublin Core där det informeras om DC:s uppkomst och utveckling, olika DC-publikationer och projekt som använder sig av DC etc. Bl.a. står att läsa följande om målsättningen:

- ”simplicity of creation and maintenance
- commonly understood semantics
- international scope and applicability
- extensibility

<sup>31</sup> <URL: <http://linnae.helsinki.fi/meta/nmfinal.htm>> 981130

<sup>32</sup> Se ordlistan.

<sup>33</sup> Från engelskans ”Dublin Core Directorate”.

<sup>34</sup> <URL: <http://wwwub2.lu.se/tk/warwick.html>> 981030



- interoperability among collections and indexing systems.”<sup>35</sup>

Även om Dublin Core syftar till att vara så enkelt att i princip vem som helst ska kunna använda det, måste det, eftersom Dublin Core är något som bibliotek och museer intresserar sig för, även finnas utrymme för att så exakt och utförligt som möjligt kunna beskriva dokumentobjekten. Den katalogisering som Dublin Core utgör ska också gå att förena med den katalogisering som redan existerar inom dessa organisationer. T.ex. finns det program som kan mappa<sup>36</sup> DC-metadata till MARC-katalogisering och tvärtom.

Dublin Core består av femton beskrivningselement eller kärnelement (core element) såsom Subject, Creator, Publisher etc. Elementen är upprepningsbara, d.v.s. man kan t.ex. lägga in antingen ett eller fem Subject-fält. Inom dessa femton element finns sedan tre s.k. qualifiers (ung. bestämningsord). Med hjälp av dessa kan man specificera informationen i ett element. På så sätt finns det både en enkelhet i utformningen samtidigt som det finns utrymme för specialisering. Dock finns det en konflikt mellan dessa begrepp; hur enkelt eller hur specialiserat det bör vara. Juha Hakala beskriver hur det inom DC-kretsar finns en viss spänning mellan s.k. minimalister och strukturalister<sup>37</sup>. Minimalisterna vill inte ha några qualifiers alls - allt för att det ska vara så enkelt som möjligt - medan strukturalisterna vill ha möjlighet att kunna göra utförligare resursbeskrivningar och vill därmed ha qualifiers vilka ger större finfördelning och precision inom respektive element. Projekt som arbetar med elektronisk publicering kan utifrån sina behov skapa egna varianter av dessa element. De kan ta bort element de tycker är onödiga, skapa egna qualifiers o.s.v. På så sätt är DC ett flexibelt system. Det föreligger dock en viss risk med detta förfarande. Om varianterna blir alltför individuellt utformade faller tanken på DC-formatet som standard. Program och sökrobotar kan då få problem med att hantera metadata-informationen korrekt.

Vissa DC-element har tillhörande kontrollerade vokabulärer som visar vad man kan välja att lägga in i elementen, t.ex. vilka sorters dokumenttyper man kan välja att lägga in i Type-elementet. Detta dels för att förhindra att metadataproducenterna lägger in ord och begrepp som ingen annan än de själva begriper, men framförallt för att underlätta för programvaran i söktjänsterna som ska avläsa metadatan. Den inlagda metadatan ska fungera i olika samlingar och programsystem, tvärs över programgränserna. Flera kontrollerade vokabulärer finns redan och andra är under utveckling. Ansvariga för DC:s utveckling rekommenderar att man använder befintliga vokabulärer.

Nedan följer en sammanfattande redogörelse för de femton DC-elementen. Vissa av dem är fortfarande under utveckling och det är inte helt klart hur de i framtiden kommer att användas<sup>38</sup>.

1. Title (Label: titel): Det namn författaren/författarna eller utgivaren väljer åt dokumentet.

<sup>35</sup> Dublin Core Metadata Initiative. <URL: [http://purl.oclc.org/metadata/dublin\\_core/main.html](http://purl.oclc.org/metadata/dublin_core/main.html)> 981130

<sup>36</sup> Se ordlistan.

<sup>37</sup> Hakala, Juha. *Dublin Core in 1997 a report from Dublin Core metadata workshops 4 & 5.*

<URL: <http://linnea.helsinki.fi/meta/dcnord.html>> 981030

<sup>38</sup> För mer information om elementen se <URL: [http://purl.org/dc/about/element\\_set.htm](http://purl.org/dc/about/element_set.htm)>

2. Author or Creator (Label: Creator): Person eller organisation som ansvarar för skapandet av resursens intellektuella innehåll. Författare när det gäller skrivna dokument, fotograf när det gäller bild etc.

3. Subject and Keywords (Label: Subject): Kontrollerade vokabulärer såväl som fria nyckelord vilka beskriver resursens ämne eller innehåll. Detta element är oftast mycket utförligt med tillhörande kontrollerade vokabulärer och ämnessystem ur vilka man väljer beskrivningsord. Det är viktigt att man inte använder ord slumpmässigt.

Ämnesordslistorna gör att resursbeskrivningarna blir så enhetliga som möjligt. Man kan även använda fria nyckelord som man tycker bäst beskriver dokumentets innehåll och med sådana ord man tror sökes på.

4. Description (Label: Description): En kortare beskrivning av resursens innehåll, ofta en sammanfattning eller en abstract.

5. Publisher (Label: Publisher): Enheten som ansvarar för att sidan ges ut, t.ex. ett förlag, universitet eller företag.

6. Other Contributor (Label: Contributor): Någon som i relativt hög grad medverkat till resursens intellektuella innehåll, men ej i lika hög grad som författaren, t.ex. medförfattare, illustratör eller handledare.

7. Date (Label: Date): Tidsangivelse för diverse företeelser: Oftast ”sidans skapande”, ”tillgängliggörande” eller ”bäst-före-datum”.

8. Resource Type (Label: Type): Vilken kategori resursen kan räknas till, t.ex. hemsida, roman, essä, teknisk rapport, uppslagsverk etc. Detta element är fortfarande under utveckling.

9. Format (Label: Format): För att kunna identifiera resursens dataformat, mjukvarans och eventuellt också hårdvarans, i syfte att kunna använda resursen. Detta element är fortfarande under utveckling.

10. Resource Identifier (Label: Identifier): Ett sätt att identifiera resursen, att tala om var resursen kan återfinnas, t.ex. URL, URN eller ISBN. Det är ofta meningen att man ska använda URL.

11. Source (Label: Source): Information om källan bakom resursen, oftast en URL.

12. Language (Label: Language): Språket på resursens intellektuella innehåll.

13. Relation (Label: Relation): Här kan man länka till ”besläktade” resurser, eller visa att det är en del av en serie t.ex. Detta element är fortfarande under utveckling.

14. Coverage (Label: Coverage): Resursens geografiska eller tidsmässiga täckning, d.v.s. vilken plats eller tidsperiod som är aktuell.

15. Rights Management (Label: Rights): Rättigheter, oftast copyright. Detta element är fortfarande under utveckling.

För skapande av DC-metadatan finns vissa hjälpfunktioner såsom inmatningsformulär, användarguider, stödfunktioner etc. som ska göra det enkelt för alla användare att lägga in korrekt metadata i sina dokument. Den kanske viktigaste av dessa hjälpfunktioner är inmatningsformuläret. Detta genererar DC-metadatan. I ett sådant nätbaserat formulär ligger alla elementen som fält, vilka dokumentproducenten fyller i med uppgifter om sitt dokument. Uppgifterna omvandlas sedan till HTML-kodade metataggar<sup>39</sup>, vilka man klistrar in i det aktuella dokumentet, mellan <HEAD> och </HEAD>. Då har s.k. inbäddad metadata blivit skapad, d.v.s. metadata som rent fysiskt hamnar i det dokument som beskrivs. En metatagg kan se ut så här:

```
<META NAME=DC.Subject" CONTENT="lövgrodor">
```

Denna tagg beskriver ett dokument som innehåller ämnet "lövgrodor". Själva "META NAME" i HTML-koden är standard vad gäller metadata i allmänhet. Vad som visar att det är just DC-metadata i den här taggen är bokstäverna "DC" framför "Subject".

För att få en uppfattning om hur ett formulär ser ut, kan man närmare studera "Nordic DC metadata creator"<sup>40</sup>. Detta formulär, skapat inom ramarna för Nordiska metadataprojektet, är det formulär vi använt i vår jämförelse p.g.a. dess generalitet.

Man kan också förvara metadata på ett annat sätt, nämligen som separat katalogpost. Metadata-uppgifterna existerar då utanför dokumentet och blir ungefär som ett elektroniskt katalogkort, och kan förvaras i metadata-databaser, såsom t.ex. Nordiskt Web Index<sup>41</sup>, för närvarande det enda helt "metadatakänsliga" indexet. I en sådan metadata-databas görs sökningarna enbart i metadata-informationen där en URL pekar mot det aktuella dokumentet. På så sätt ligger dokumenten kvar på de servrar<sup>42</sup> där dokumentproducenterna har placerat dem, och sökmotorn behöver inte hämta hela dokumenten, eller ens söka i dem för att kunna referera till dem. Det innebär också att man kan skapa en metadatapost om ett dokument utan att ha tillgång till den server där dokumentet ligger.

I rapporten från Nordiska metadataprojektet beskriver Hakala m.fl. projektets syfte att utveckla olika DC-verktyg<sup>43</sup>. Inblandade i projektet var metadata-intressenter från de skandinaviska länderna, bl.a. NetLab. Projektet var från början finansierat av NORDINFO<sup>44</sup>. Författarna konstaterar att liksom själva Dublin Core har nu (juli 1998) DC-verktygen en solid struktur som inte kommer att förändras så radikalt. Verktygen är allmänt ganska kända och används av t.ex. Nordiskt Web Index och SAFARI. Andra verktyg, som här endast nämns, är omformaren av

---

<sup>39</sup> Se ordlistan.

<sup>40</sup> <URL: <http://www.lub.lu.se/cgi-bin/nmdc.pl>> 990410

<sup>41</sup> <URL: <http://nwi.ub2.lu.se/?lang=sv>> 990410

<sup>42</sup> Se ordlistan.

<sup>43</sup> <URL: <http://linnae.helsinki.fi/meta/nmfinal.htm>> 981130

<sup>44</sup> Se ordlistan.

Dublin Core till MARC-format<sup>45</sup>, ett metadata insamlar- och indexeringsprogram<sup>46</sup> och en URN-generator (för mer information om dessa, se Nordiska metadataprojektet<sup>47</sup>). Hakala gör uppskattningen att det kanske tar två år innan DC är helt stabilt och innan dess blir det antagligen nödvändigt att revidera ”verktygslådan”<sup>48</sup>.

Vad beträffar användningen av Dublin Core är formatet fortfarande inte utbrett i de stora söktjänsterna. DC används mest i speciella ämneslänksamlingar eller kvalitetssöktjänster såsom de vi undersökt. Vissa av de stora söktjänsterna, t.ex. AltaVista är känsliga för metadata i allmänhet (inte bara DC) men använder sig bara av Keywords och Description.

På DC:s hemsida finns beskrivet ett 30-tal projekt som använder sig av Dublin Core. Projekt från Skandinavien, Europa, USA och Australien, varav de flesta är baserade i Europa. Ett projekt värt att nämnas i sammanhanget är BIBLINK, ett EU-projekt inom bl.a. UKOLN, vilket syftar till att skapa ett samarbete mellan nationella bibliotek och förläggare av elektroniskt material. BIBLINK har skapat en DC-variant, den s.k. BIBLINK Core, som gör det möjligt för förläggare av elektroniskt material att skicka in resursbeskrivningar som är tillräckliga för den nationella katalogiseringen<sup>49</sup>. Detta visar återigen hur flexibelt DC är. Det kan användas från en enkel webbdokumentbeskrivning av Svea Svenssons hemsida till att skapa en avancerad nationell bibliografisk post som beskriver Svea Svenssons avhandling om lövgrodors matvanor och allmänna leverne.

Speciellt i Skandinavien är Dublin Core ganska utbrett. I Danmark och Finland har Dublin Core blivit valt till det metadata-format som ska användas för resursbeskrivning av officiella elektroniska dokument från statliga och kommunala organ.

Vad beträffar reglerna gällande inmatning av DC-metadata, påpekar Hakala att det inte finns några utarbetade katalogiseringsregler. I viss mån fungerar användarguiderna som sådana, men dessa är inte tillräckliga för allas behov. Därför bör varje seriöst DC-projekt utforma sina egna användarguider som dock inte bör variera allt för mycket från den generella guiden.

#### 4.2.1 Qualifiers

Förutom de femton elementen finns det ytterligare utrymme att specificera en resursbeskrivning med s.k. qualifiers. Man kan inte lägga till fler element än de redan existerande, men man kan specificera ett element genom att lägga till en qualifier i elementet. Vad en qualifier gör är att den ”snävar ner” innehållet i elementet på ett standardiserat vis. Kontrollerade vokabulärer återfinns som qualifiers, och fler sådana är under utveckling. Det finns tre qualifiers; Type, Language och Scheme, varav Type och Scheme är vanligast förekommande.

<sup>45</sup> Från engelskans ”Dublin Core to MARC converter”.

<sup>46</sup> Från engelskans ”metadata harvesting and indexing application”.

<sup>47</sup> *The Nordic metadata project Final report*. Kapitel 2 ”The metadata toolbox.”

<URL: <http://linna.helsinki.fi/meta/nmfinal.htm>> 981130

<sup>48</sup> Ibid.

<sup>49</sup> <URL: <http://hosted.ukoln.ac.uk/biblink/>> 981030

**Type** är tänkt att begränsa det språkliga innehållet i elementet, d.v.s. ta ett steg ner i ”betydelsehierarkin”. Därför kallas också type-qualifiern ”subelement” eftersom det handlar om ett underelement (fungerar ungefär som i SAB-klassningen, där t.ex. avdelning Ughf är en underavdelning till Ug). Som exempel tar vi Svea Svenssons avhandling om lövgrödor.

Exempel på metatagg i Title-elementet:

```
<META NAME="DC.Title" CONTENT="Lövgrödors matvanor och allmänna leverne.">
```

Exempel på metatagg i Creator-elementet:

```
<META NAME="DC.Creator.PersonalName"CONTENT="Svensson, Svea">
```

I det senare exemplet finns en type-qualifier; ”PersonalName”. Det specificerar kärnelementet ”Creator”. I nämnda exempel beskrivs dokumentet där upphovsmannens namn är Svea Svensson. Det kan finnas andra variationer av Creator än ett personligt namn, om en organisation eller en förening är upphovet till ett dokument. Då blir type-qualifiern ”CorporateName”. Exempel:

```
<META NAME="DC.Creator.CorporateName"CONTENT="Svenska Lövgrodeföreningen">
```

Om föreningen har en e-postadress kan detta också ligga som en type-qualifier i Creator.

Exempel:

```
<META NAME="DC.Creator.CorporateName.Adress"CONTENT="sv_lov@telia.com">
```

**Language** specificerar språket i elementinnehållet. Om ett dokument är skrivet på engelska visas det i elementet Language, men med en language-qualifier kan man specificera vilket språk som används i taggarna. T.ex. om det finns ord på engelska i Subject (och det i övrigt är ett svenskt dokument) kan man specificera detta genom att lägga in en language qualifier. Exempel:

```
<META NAME="DC.Subject"LANG="eng"CONTENT="frogs">
```

**Scheme** är ett bestämningsord som påvisar en kodstandard eller dylikt t.ex. SAB eller UDK<sup>50</sup>.

Exempel:

```
<META NAME="DC.Subject"SCHEME="SAB"CONTENT=Ughf">
```

Ett annat exempel är Date-elementet där man måste välja ett scheme som visar hur datum ska skrivas. Standarden är ”ISO8601” d.v.s. datum skrivs ut så här: 1999-02-10 för 10 februari 1999. Exempel:

```
<META NAME="DC.Date"SCHEME="ISO8601"CONTENT="1999-02-10">
```

---

<sup>50</sup> Se ordlistan.

## 4.2.2 RDF och The Warwick Framework

RDF, Resource Description Framework, och Warwick Framework är termer som återkommer i många artiklar om metadata och Dublin Core. Termerna hör hemma i resonemanget om hur metadata rent tekniskt fungerar vad beträffar programvaror etc. Ett kortfattat sätt att förklara RDF är att kalla det en typ av språk för metadata:

”A language for expressing metadata which is simple to process by machine.”<sup>51</sup>

RDF växte fram ur ett behov av en generell kodning av metadata. RDF, som fortfarande är under utveckling, har utarbetats genom ett samarbete mellan olika metadatagrupper och W3C Metadata Project<sup>52</sup> i syfte att stödja nyttjandet av metadata på webben. Därför rekommenderar W3C RDF som en standard för hur man bör koda metadata så att den blir ”maskinbegriplig”. RDF är en sorts ram inom vilken olika metadataformat fungerar ihop. Dublin Core är ett av dessa format. Just nu pågår granskning av varje element i Dublin Core i syfte att dessa ska passa in i RDF-standarden. Om c:a ett år kommer Dublin Core att rekommendera RDF. Vad gäller Warwick Framework är det ett försök till strukturering av s.k. metadata-behållare vilka gör det möjligt för separat metadata att existera utanför dokumentet. Dublin Core är ett av de olika metadataformat som använder sig av Warwick Framework. MARC-formatet är exempel på ett annat format. Om detta skriver Rasmus Bruun i sin magisteruppsats ”Er lykken gjort?...” där han förklarar Warwick Framework sålunda<sup>53</sup>:

”Med en container åbnes der muligheder for at et dokument's metadata ikke befinder sig i selve dokumentet, men i et separat dokument med en angivelse av det beskrevne dokument's URL ... Denne struktur gør det muligt at tildele et dokument metadata i overensstemmelse med flere forskellige metadataformater på en overskuelig måde.”<sup>54</sup>

## 4.2.3 Dublin Core Workshops

Dublin Core-formatet har utvecklats under ett antal workshops. Deltagarna är speciellt inbjudna och kommer från skilda om än besläktade yrkeskategorier såsom systemutvecklare, bibliotekarier, databasexperter, katalogisatörer, museipersonal, arkivarier m.fl. Mellan femtio och åttio personer har deltagit varje gång. Här följer en kortfattad genomgång av dessa workshops.

<sup>51</sup> Swick, Ralph. *W3C Technology and Society domain, Metadata Activity*. 1998.

<URL: <http://www.w3.org/Metadata/Activity.html>> 981130

<sup>52</sup> W3C Metadata Home Page: <URL: <http://www.w3.org/Metadata>>981130. Förklaring av W3C, se ordlistan.

<sup>53</sup> Bruun, Rasmus. *Er lykken gjort? - om metadata som indekseringsgrundlag ved bibliografisk kontrol af webdokumenter*. Danmarks Biblioteksskole. 1997.

<URL: <http://www.db.dk/student/k96/4/rb/opgaver/2semopg%5Ftekst.html>>981130

<sup>54</sup> Ibid.

### **DC-1: Mars 1995 i Dublin, Ohio, USA<sup>55</sup>**

Organiserat av Online Computer Library Center Office of Research (OCLC) och The National Center for Supercomputing Applications (NCSA).

På detta första möte ville man komma fram till ett enkelt och praktiskt sätt att beskriva elektroniska dokument. Huvudsyftet med detta var att underlätta återvinning av resurser på Internet. Deltagarna satte som mål att utveckla ett kärnset av beskrivningselement för dokument-liknande objekt samt att undersöka hur man skulle kunna beskriva även andra typer av objekt, såsom ljudfiler, videoclips, bilder o.a.

Syftet med projektet var att tillmötesgå informationssökare och ge dem möjligheter att på ett funktionellt, tidsbesparande och kostnadseffektivt sätt hitta fram till elektronisk information. Vid denna tidpunkt ansågs mycket material av elektronisk art i det närmaste omöjlig att lokalisera. Om dokument skulle kunna hittas genom systematiskt sökande, måste de beskrivas på ett sätt som motsvarar deras intellektuella innehåll. Detta skulle kunna göras genom en generator, som kunde kopplas till en databas med pekare till dessa dokument. Beskrivningen skulle innehålla mer information än en traditionell indexingång, men mindre än en formell katalogpost. Elektroniska verktyg för hjälp av återsökande av elementen och insamlande av dem i databaser skulle kunna utvecklas. Mötesdeltagarna var överens om att det bästa sättet var att uppmuntra webbprocenterna att beskriva sina dokument och att beskrivningselementen skulle utformas på ett sådant sätt att ingen formell utbildning skulle krävas. Att utbildade katalogisatörer skulle sköta handhavandet av beskrivningarna ansågs orimligt. Det skulle bli alltför kostsamt och tidsödande.

DC-1 resulterade i tretton metadata-element som kunde användas för att beskriva elektroniska resurser. Dessa kallades the Dublin Metadata Core Element Set. Semantiken hos elementen ansågs så pass elementär att många olika typer av användare skulle kunna förstå den. Elementen skulle också kunna mappas över till andra system, såsom USMARC<sup>56</sup>. Alla element var frivilliga och upprepningsbara. Utrymme för vidare utveckling hos elementen gavs. Elementen var: Title, Creator, Subject, Description, Publisher, Contributor, Date, Identifier, Source, Language, Relation, Coverage och Rights.

Qualifiers diskuterades också, men man kom inte fram till någon gemensam standard vad beträffar dessa. Det ansågs viktigt att upprätthålla en kontinuerlig kommunikation mellan de olika enheterna som ingått i denna första workshop.

### **DC-2: April 1996 i Warwick, England<sup>57</sup>**

Organiserat av OCLC och UK Office of Library Networking (UKOLN).

Tre huvudfrågor diskuterades på detta andra möte:

- Att utarbeta en konkret syntax för DC uttryckt i SGML<sup>58</sup>. Detta skulle sedan mappas till HTML.

<sup>55</sup><URL: <http://purl.org/dc/workshops/dc1conference/index.htm>> 981230

<sup>56</sup> Se ordlistan.

<sup>57</sup><URL: <http://www.oclc.org:5046/oclc/research/conferences/metadata2/>> 981230

<sup>58</sup> Se ordlistan.

- The Warwick Framework.
- Att utarbeta instruktioner för skapande och bevarande av metadata. En guide till genererande av resursbeskrivningar och en guide till administratörer av samlingar.

### **DC-3: September 1996 i Dublin, Ohio, USA<sup>59</sup>**

Organiserat av OCLC och Coalition for Networked Information (CNI).

Det huvudsakliga syftet med denna tredje workshop var att utarbeta ett sätt att beskriva bildobjekt. Många av deltagarna var inställda på att ett separat set med beskrivningselement just för bilder var den enda lösningen. Det visade sig dock att det räckte med att utöka de tretton ordinarie elementen till femton. De två nya täckte behovet av bildbeskrivning. Flera av deltagarna blev glatt överraskade av detta faktum. Så här uttrycker sig Charles Rhyne som deltog i DC-3:

“I was especially surprised that we concluded that the elements needed to discover text and images on the internet are similar. The text and the images themselves are radically different and require different types of expertise to study and interpret them, but most of the primary categories under which we classify and search for them are similar.”<sup>60</sup>

De två nya beskrivningselementen med tillhörande exempel var följande:

Typ: t.ex. text, bild, video.

Format: t.ex. TIFF<sup>61</sup>, GIF<sup>62</sup>, PICT<sup>63</sup>.

### **DC-4: Mars 1997 i Canberra, Australien<sup>64</sup>**

Organiserat av OCLC, the Distributed Systems Technology Centre (DSTC) och the National Library of Australia (NLA).

Denna workshop kom att bli ett möte mellan s.k. minimalister och strukturalister. Minimalisterna ville ha en så enkel DC-uppsättning som möjligt. Dels för att underlätta för metadataproducenten, dels för att slippa de kostnader det skulle innebära att låta utbildade katalogisatörer sköta metadata-inmatningen (vilket skulle bli följden om inmatningen blev för komplex) och dels för att främja interoperabilitet. Minimalisterna ville att elementen skulle betyda samma sak överallt och enkelhet skulle underlätta detta syfte. Man menade att ett enkelt system skulle underlätta både för användarna och för elektroniska verktyg som inte kunde urskilja alltför detaljerad metadata. Det negativa med minimalisternas tankegång var att sökresultaten riskerade att bli många och med låg precision. Strukturalisterna å andra sidan ville ha många valmöjligheter vad beträffar resursbeskrivningen. Detta skulle dock medföra nödvändig och dyr utbildning av katalogisatörer, men å andra sidan hög precision vad beträffar sökresultat. Man ville också

<sup>59</sup> <URL: <http://www.oclc.org:5046/oclc/research/conferences/imagemeta/>> 981230

<sup>60</sup> Rhyne, Charles. Chair of Art History at Reed College.

<URL: <http://www.dlib.org/dlib/january97/01weibel.html>> 981130

<sup>61</sup> Se ordlistan.

<sup>62</sup> Se ordlistan.

<sup>63</sup> Se ordlistan.

<sup>64</sup> <URL: <http://www.dstc.edu.au/DC4/>> 981230



göra det möjligt att anpassa elementen till speciella ändamål och individuella projekt. Målet var att DC skulle bli en sorts gyllene medelväg mellan minimalister och strukturalister.

Tre huvudfrågor diskuterades:

- Utvecklingsfrågor för att stabilisera DC.
- Elementstruktur med bestämmande underavdelningar, s.k. qualifiers. På tidigare workshops hade införandet av qualifiers diskuterats. Nu utarbetades the Canberra Qualifiers. Dessa var:
  - Language
  - Scheme
  - Type
- Klarare definitioner för vissa av elementen. Främst Coverage, Relation och Rights.

Ett annat område som diskuterades var hur den framarbetade metadatamodellen skulle uttryckas i HTML. Hur skulle metadata kodas och transporteras? Diskussioner kring bl.a. XML<sup>65</sup> och PICS<sup>66</sup> förekom.

#### **DC-5: Oktober 1997 i Helsingfors, Finland<sup>67</sup>**

Organiserat av OCLC, National Library of Finland och the Coalition for Networked Information (CNI).

Vid den här tidpunkten hade flera projekt som tillämpade metadata startats i minst tio länder runt om i världen. DC hade rönt framgång och såg ut att kunna accepteras som en standard. Särskilt stor uppmärksamhet hade formatet fått i Australien och Danmark.

Tekniska lösningar vad beträffar inbäddad metadata i HTML 2.0 och HTML 4.0 diskuterades på DC-5. Vidare rörde diskussionerna en datumstandard och deltagarna kom fram till att applicera ISO 8601 i metadatamodellen. RDF-utveckling stod också på dagordningen. De tre DC-elementen Date, Coverage och Relation diskuterades ingående. På DC-5 arbetade man med att fastslå de femton elementens semantiska betydelse. Man ansåg nu elementen relativt stabila, naturligtvis med öppenhet för ytterligare utveckling. Man kallade stabiliseringen ”the Finnish Finish”.

#### **DC-6: Höst 1998 i Washington, USA<sup>68</sup>**

Denna senaste workshop ägde rum i november 1998, alltså mitt under pågående uppsatsskrivande för oss. På DC:s hemsida hade några veckor före uppsatsens slutförande information om mötet ännu inte lagts ut. Vi har dock sammanfattat de punkter som fanns på dagordningen för mötet samt snappat upp en del information genom våra handledare på NetLab i Lund.

På DC-6 koncentrerade man sig på att slutföra ”the Finnish Finish”. Nu ansågs det bakomliggande språkbruket vara relativt fastslaget, naturligtvis med plats för mindre ändringar.

---

<sup>65</sup> Se ordlistan.

<sup>66</sup> Se ordlistan.

<sup>67</sup> <URL: <http://linnea.helsinki.fi/meta/DC5.html>> 981230

<sup>68</sup> <URL: <http://purl.org/dc/workshops/dc6conference/index.htm>> 981230

De femton elementen skulle heller inte utökas, utan förbli konstanta. Denna fastslagna standard skulle få namnet DC 1.0. Man diskuterade också metadata i förhållande till RDF och XML. The DC Data Model Working Group rapporterade om hur de arbetar med omvandling av DC-metadata till RDF. Man började också förbereda ett omarbetat DC, med utrymme för hantering av rättigheter i samarbete med musikindustrin. Diskussioner om hur DC-elementen ska göras sökbara genom Z39.50<sup>69</sup> föregick också. Särskilt när det gäller qualifiers.

### 4.3 SAMORDNINGSFÖRSÖK

Sedan drygt ett år tillbaka har inblandade i DC-projekt runt om i landet diskuterat huruvida det skulle vara möjligt att samsöka i de olika söktjänster som tillämpar DC. Syftet med DC som standard är att den ska kunna fungera i ett större sammanhang i framtiden. I dagsläget är standarden så pass utvecklad att den fungerar i mindre databaser och tiden verkar nu mogen för samarbete över projektgränserna. För närvarande pågår samarbete mellan söktjänsterna SAFARI, Svenska miljönätet och Svesök. Syftet med samarbetet är att anpassa de inbördes söktjänsternas sätt att tillämpa metadata och därigenom underlätta eventuella framtida samsökningar i dem.

Åsa Präntare på Svenska miljönätet har jämfört SAFARI:s och Svenska miljönätets sätt att tillämpa DC-formatet samt kommit med förslag på lösningar som ska passa de båda söktjänsterna<sup>70</sup>. Sten Hedberg på Uppsala universitetsbibliotek är en av nyckelpersonerna i diskussionerna kring en gemensam standard för söktjänsterna. Han har i detalj jämfört och sammanställt de tre söktjänsternas inbördes tillämpning av formatet<sup>71</sup>. Sten Hedberg arbetar genom nätprojektet "Nya vägar för boken" bl.a. med att försöka utforma en svensk DC-praxis<sup>72</sup>.

### 4.4 Z39.50

Samordningen ska realiseras genom Z39.50-standard. Detta är ett amerikanskt nätverksprotokoll<sup>73</sup> som är accepterat av ISO<sup>74</sup> och som gör det möjligt för två datasystem att utbyta information om både klient<sup>75</sup> och server har Z39.50-funktion<sup>76</sup>. Utvecklingsprocessen kring protokollet har pågått sedan tidigt 80-tal. Från början var syftet med Z39.50 att underlätta kommunikation och överföring av bibliografiska poster i MARC-

---

<sup>69</sup> Se ordlistan.

<sup>70</sup> Präntare, Åsa. Jämförelse mellan SAFARI:s och Svenska miljönätets tillämpning av DC-metadata, december 1998.

<sup>71</sup> Hedberg, Sten. Jämförelse mellan SAFARI:s, Svenska miljönätets och Svesöks tillämpning av DC-metadata, december 1998.

<sup>72</sup> Hedberg, Sten. *Bruket av metadata enligt Dublin Core: Principer, teknik och tillämpningar utanför Sverige*. <URL: <http://www.kb.se/nvb/Metadata/meta1.htm>> 990430

<sup>73</sup> Se ordlistan.

<sup>74</sup> Se ordlistan.

<sup>75</sup> Se ordlistan.

<sup>76</sup> Turner, Fay. "Selecting a Z39.50 Client or Web Gateway". *Library Hi Tech*. Nr 16(2) 1998, s 7-18.

formatet mellan olika bibliotekssystem eller databaser. Det blev dock något mycket mer av denna tidiga tanke. Fredrik Hammar som skrivit en magisteruppsats om Z39.50-protokollet beskriver utvecklingsprocessen:

“Det framgick ganska snart under utvecklingen att det hade blivit ett mer allmänt och generellt hållet protokoll än vad man först hade eftersträvat. Antalet användningsområden blev betydligt fler än väntat och medförde att det nu kunde användas till mer allmänna informationssökningar via Internet/www och flertalet bibliotekssystem (t.ex. Horizon SLS/Libertas) och även byggas in i dessa.”<sup>77</sup>

Den senaste versionen (version tre) kom 1995 och godkändes av ISO i mars 1997. Den tekniska utvecklingen av protokollet står Z39.50 Implementors Group (ZIG) för<sup>78</sup>.

Hur fungerar då Z39.50-protokollet? Jo, det finns många informationssystem i form av databaser och de är alla individuellt strukturerade och serverar användaren olika utformade gränssnitt. Detta skapar naturligtvis problem för informationssökaren, eftersom det är energikrävande och tidsödande att lära sig flera olika system. Med hjälp av Z39.50 kan dessa system kopplas samman och användas samtidigt, utan att användaren behöver komma i kontakt med något annat gränssnitt än det han/hon är van vid. Man kommer alltså åt en ansevärt större mängd information genom ett välbekant och förhoppningsvis lättfattligt gränssnitt. Syftet med att ta fram detta hjälpmedel är alltså att underlätta för informationssökaren.

Z39.50 har ännu inte till fullo slagit igenom bland bibliotekarier och informationsexperter. Det vinner dock i allt högre utsträckning mark bland dessa yrkesgrupper. Speciellt i USA, men även i Europa har bibliotek testat Z39.50<sup>79</sup>. Andrew Cranfield har beskrivit Z39.50 i den danska tidskriften *DF-revy*:

“Der er ingen tvivl om, at Z39.50 vil blive en integreret del av fremtiden, når vi taler om biblioteks- og informationssystemer. Der obnes op for en helt ny verden, i det vi kan foretage bibliografiske søgninger, genbruge katalogiseringer, bestille og evt. betale for dokumenter fra andre biblioteker fra den samme og velkendte grænseflade. Det er ikke længere nødvendigt for brugeren at mestre ny viden om hver enkelt database - man vill i princippet kunne søge i f.eks. DanBib, Bibsys, LIBRIS, Blaise, OCLC o.s.v. som om den var en og samme base og i samme session - hvilken herlig tanke!”<sup>80</sup>

Detta citat speglar den nu rådande synen inom biblioteks- och informationsvärlden vad beträffar Z39.50-protokollet.

<sup>77</sup> Hammar, Fredrik. *Nätverksbaserade informationssystem: Standardprotokoll ANSI Z39.50/OSI SR*. Högskolan i Borås. 1997, s.16.

<sup>78</sup> <URL: [www.ukoln.ac.uk/dlis/z3950/defin.html](http://www.ukoln.ac.uk/dlis/z3950/defin.html)> 981230

<sup>79</sup> Hammar, s. 23f.

<sup>80</sup> Cranfield, Andrew. ”Z39.50 og danske biblioteker - en temadag: Danmarks Biblioteksskole, onsdag d. 3. december 1997.” *DF-revy*. 21(2) 1998, s. 47f.

Rent praktiskt går det till på följande sätt: Informationssökaren gör en sökning i en databas med ett för honom/henne välbekant gränssnitt. Frågan skickas till klienten som översätter frågan till Z39.50-format och i sin tur skickar vidare frågan till servern. Servern översätter sedan frågan i Z39.50-format till det språk som mottagardatabasen talar. När så frågan slutligen har besvarats skickas den tillbaka via servern och klienten till informationssökaren, som får svaret via samma gränssnitt som han/hon använt sig av hela tiden.

För att avrunda detta kapitel om Z39.50 vill vi återigen låta Fredrik Hammar formulera vikten av detta interoperabilitetsverktyg:

“Sammanfattningsvis kan det sägas att såväl systemleverantörer som bibliotek ser i Z39.50-protokollet en stor utvecklingspotential som kan bidra till att lösa flera av bibliotekens informationsproblem och flaskhalsar. Informationssökningsprocessen kommer att revolutioneras. Protokollet kommer att förändra såväl bibliotekens traditionella informationsstrukturer som slutanvändarnas informationsbehov/vanor.”<sup>81</sup>

#### 4.5 DUBLIN CORE I FRAMTIDEN

Carl Lagoze skriver i artikeln ”The Warwick Framework” om DC-2 då Dublin Core var ett år gammalt. Där ställde man sig vissa frågor, bl.a. om man verkligen kunde förvänta sig att webbproducenter skulle lägga in metadata på sina dokument<sup>82</sup>. Detta är ett ämne som framtiden får utvisa. Fungerar idén bakom DC, d.v.s. att det är dokumentproducenterna själva som ska lägga in metadata på sina dokument? Är det realistiskt att kräva av människor att de ska lägga in metadata själva? Kommer de verkligen att göra det i någon större utsträckning?

I sin artikel ”The Dublin Core and Warwick Framework...” skriver Thiele om en annan aspekt, nämligen vilken påverkan metadata kommer att ha på Internet. Han efterlyser rapporter om hur effektivt DC är jämfört med andra metadataformat då den redan existerande litteraturen kring DC mestadels handlar om beskrivning av själva Dublin Core, hur det kom till och hur det fungerar etc. Han efterlyser även rapporter om hur den tekniska utvecklingen ser ut i sammanhanget, samt vilka större vidgående effekter metadata får på Internet. Thiele frågar sig om inte denna form av indexering kommer att dela upp resurserna på Internet i två tydliga grupper. En grupp av resurser som är associerade med traditionella akademikerkringar och forskarmiljöer där man använder sig av resursbeskrivning och en annan grupp som står utanför den akademiska världen där man inte kommer att använda sig av resursbeskrivningar. Blir då metadata en slags kvalitetsstämpel, frågar sig författaren. Något som bestyrker värdet av dokumentet<sup>83</sup>?

---

<sup>81</sup> Hammar, s. 62.

<sup>82</sup> Lagoze, Carl. ”The Warwick Framework - A Container Architecture for Diverse Sets of Metadata.” *D-Lib Magazine*. Juli/Augusti 1996. <URL: <http://www.dlib.org/dlib/july96/lagoze/07lagoze.html>> 981030

<sup>83</sup> Thiele, Harold. ”The Dublin Core and Warwick Framework A Review of the Literature, March 1995 - September 1997.” *D-Lib Magazine*. Januari 1998. <URL: <<http://mirrored.ukoln.ac.uk/lis-journals/dlib/dlib/dlib/january98/01thiele.html>> 981030

## 5. PRESENTATION AV SÖKTJÄNSTERNA

Vi har valt att undersöka de tre kvalitetssöktjänsterna SAFARI, Svenska miljönätet och Svesök. Nedan följer en presentation av projektens innehåll, målsättning och kvalitetskriterier samt en översiktlig beskrivning av vilka fält som används i respektive söktjänsts metadataformulär. Kapitlet bygger på projektbeskrivningar på nätet samt intervjuer med projektansvariga.

### 5.1 SAFARI<sup>84</sup>

#### 5.1.1 Innehåll och målsättning

Namnet SAFARI står för ”Spridning av forskningsinformation till allmänheten över Internet”. Högskoleverket ansvarar för projektet, vilket är baserat på ett regeringsuppdrag. Uppdraget gick ut på att samordna ett system för svensk forskningsinformation på Internet. Projektet startade 1996, med målsättningen att de inblandade databaserna skulle fungera i slutet av 1998. Detta mål är uppfyllt.

De som bidrar med forskningsinformation till SAFARI är olika forskningsenheter landet runt. De tänkta målgrupperna är allmänheten, skola, näringsliv samt forskare. Syftet med SAFARI är att det ska vara enkelt att hitta specifik forskningsinformation var än i Sverige den bedrivs. Allt material ska vara samlat på ett ställe, så att man slipper leta upp varje enskild organisations hemsida.

Inom SAFARI finns en sökfunktion i vilken man kan söka på allt forskningsmaterial inom databasen. I anslutning till sökformuläret har man också tillgång till ett ämnesträd som fungerar som bläddringsfunktion. Ämnesträdet och sökfunktionen innehåller samma informationsmängd.

SAFARI uppmanar forskare att lägga ut sin information på nätet för att:

- “rekrytera studenter
- skapa legitimitet för akademisk verksamhet och stärka ett visst ämnes status
- väcka intresse hos potentiella finansiärer
- rekrytera personal
- underlätta studenters arbete och stärka banden mellan forskning och undervisning
- underlätta tvärvetenskapliga forskningsprojekt
- profilera sin organisation genom att framhäva vissa områden
- förstärka bilden av en professionell organisation.”<sup>85</sup>

<sup>84</sup> <URL: <http://safari.hsv.se/>> 981230

Vad beträffar ämnesklassifikation har SAFARI valt att använda sig klassifikationssystemet CERIF. CERIF står för ”Common European Research Information Format” och har tagits fram av EU. Systemet består av 369 termer på tre nivåer.

Förutom den beskrivna databasen finns också en lokal variant av SAFARI i Lund. Denna kallas SiL, SAFARI i Lund<sup>86</sup> och det är Lunds universitet som ansvarar för den. Om en sökning utförs i SAFARI omfattas även SiL av denna.

### 5.1.2 Kvalitetskriterier

Det är dokumentproducenten själv, eller den forskande myndigheten, som föreslår sina resurser samt förväntas märka dem med metadata. För att bli synlig i SAFARI måste den forskare eller den forskningsenhet som vill bidra med dokument anmäla till SAFARI var resursen finns att hämta. Detta görs i samband med metadatamärkning med hjälp av inmatningsformuläret som finns att tillgå på SAFARI:s hemsida. Metadata är alltså ett måste för att resursen ska accepteras.

De forskare som är intresserade av att bidra med resurser till SAFARI uppmanas att gå igenom sex frågor. Frågorna berör kvalitetskriterier och lyder som följer:

- “1. Hur arbetar din organisation med forskningsinformation...
2. Finns redan den hemsida du vill registrera...
3. Följer sidan våra grundkrav...
4. Vet du hur sidan ska klassificeras...
5. Är du ansvarig för sidan...
6. Vet du till vem sidan vänder sig till...”<sup>87</sup>

### 5.1.3 Inmatningsformulär

Nedan följer en översiktlig redogörelse över vilka fält SAFARI tillämpar i sitt inmatningsformulär:

DC:s fältnamn	SAFARI:s fältnamn	Kommentar
Title	Titel	
Creator	Författare/textansvarig	Listrutor samt separat fält för författarens e-postadress.

<sup>85</sup> <URL: <http://www.hsv.se/verksamhet/utredningar/safari/>> 981230

<sup>86</sup> <URL: <http://www.lu.se/lu/forskn/sil/>> 981230

<sup>87</sup> <URL: <http://safari.hsv.se/metadata/6fragor.html>> 981230

Identifier	URL	
Subject: vocabulary	Ämne: vokabulär	Klassifikation med CERIF som öppnas i eget fönster.
Subject: keywords	Ämne: nyckelord	
Description	Beskrivning	
Publisher	Utgivare	Eget fält för utgivarens e-postadress.
Date	Datum: bäst före datum	Listrutor.
Type	Typ	Typverktyg öppnas i eget fönster.
Language	Språk	Listrutor.
Coverage	Täckning	Täckningsverktyg öppnas i eget fönster. Gettys thesaurus för geografiska namn.

SAFARI har dessutom ett antal fält som kallas "oöppnade fält". Det går bra att använda dessa om man så önskar. Dessa blir dock inte sökbara i databasen, utan ifylls om man önskar få dem sökbara genom andra metadatakänsliga sökmotorer.

DC:s fältnamn	SAFARI:s fältnamn	Kommentar
Title: alternative	Alternativ titel	Ingen hjälptext.
Subject: vocabulary	Ämne: vokabulär	
Subject: klassifikation	Ämne: klassifikation	
Contributor	Medförfattare	
Date: creation	Datum: skapad	
Format	Format	
Source	Källa	
Relation	Beroenden	
Coverage	Täckning	
Rights	Rättigheter	

Förutom dessa fält har SAFARI ett eget fält som generar metadata (dock ej DC-metadata):

SAFARI:s fältnamn	Kommentar
Målgrupp	Listrutor.

Det går bra att ändra språket i inmatningsformuläret till engelska.

## 5.2 SVENSKA MILJÖNÄTET<sup>88</sup>

### 5.2.1 Innehåll och målsättning

Svenska miljönätet är även det uppbyggt efter ett regeringsförslag. Det var ett av förslagen i Miljövårdsberedningens utredning "IT i miljöarbetet" (SOU 1996:92). Regeringen föreslog i budgetpropositionen 1996 att Naturvårdverket skulle bidra till att ett miljönät skulle byggas upp enligt utredningens förslag. I december 1996 beslutade riksdagen i enlighet med propositionens förslag. I december 1997 invigdes Svenska miljönätet. Miljönätet finansieras via Naturvårdverkets budget och verksamheten styrs av ett råd, sammansatt av representanter för olika användargrupper. Det publicerade materialet är miljödata, miljöstatistik och miljöinformation som är av relevans för miljöarbetet i Sverige. Webbresurserna kommer från myndigheter, företag och organisationer som uppfyller medlemskraven som Miljönätet ställer. Det krävs alltså medlemskap av de organisationer som är intresserade av att publicera material.

Miljönätet består av Miljökatalogen samt en sökfunktion. Miljökatalogen innehåller dokument med kompletta metataggar, vilket betyder att de är ämnesklassificerade. Vissa dokument är enbart fritexttaggade, vilket innebär att de inte behöver innehålla fullständig metadata. Det kan t.ex. handla om långa listor o.dyl. De dokument som enbart förses med fritexttaggar syns inte i Miljökatalogen, men återfinns via sökfunktionen om man gör en fritextsökning. Av Miljönätets c:a 2000 dokumentär 1578 fullständigt metataggade<sup>89</sup>.

I Miljökatalogen bläddrar man sig fram till en resurs genom s.k. katalogingångar. När man gör en sökning eller bläddrar sig fram i Miljökatalogen kan man välja att titta på en detaljerad katalogpost eller det faktiska dokumentet. Katalogposten fungerar dock inte som metadata, utan helt enkelt som en extra service för informationssökaren.

Det finns även andra tjänster inom Miljönätet, såsom konferenssystemet, ett debattforum som även icke-medlemmar har rätt att delta i. Här kan man även visa länkar till sin egen webbplats som icke-medlem. Inom vissa konferenser är det även tillåtet med reklam för miljöprodukter och konsulttjänster. En annan tjänst är "Vem Gör Vad i Miljö-Sverige". Detta är en matrikel över myndigheter, organisationer och konsultföretag som sysslar med miljöfrågor. Miljönätet tillhandahåller också en e-postkatalog.

Miljönätet befinner sig på utvecklingsstadiet och provar på nya former av tjänster, t.ex. har man efter önskemål från medlemmar och användare utrett möjligheten att erbjuda miljönyheter. Ansvariga för Miljönätet är måna om att värva nya medlemmar. Detta görs genom deltagande i konferenser, mässor och seminarier. På dessa presenteras tjänsterna och vad ett eventuellt medlemskap innebär. Miljönätet kontaktar tidskrifter inom miljöområdet och tar fram underlag till artiklar samt har för avsikt att kontakta användarrepresentanter för att starta en diskussion om vad som behövs i det dagliga arbetet för miljön.

<sup>88</sup> <URL: <http://smn.environ.se/>> 981230

<sup>89</sup> December 1998.



Svenska miljö nätet värnar om dialogen med medlemmarna samt gör stora satsningar på att hänga med i den tekniska utvecklingen. Metadata är en del av denna satsning. Målsättningen är att sprida sin information till en större publik, både inom och utom Sverige, så att Miljö nätet ska bli ett reellt redskap i miljö arbetet. För närvarande deltar Miljö nätet i två projekt, med syfte att uppnå detta mål:

- “Deltar i pilotprojekt som ska studera hur Miljö nätets databas kan utgöra grunden för Sveriges rapportering till Europeiska miljö byråns (EEA) datakatalog. Metadata från Miljö nätets databas konverteras och förs över till EEA:s databas.
- Databasen kommer också anpassas så att sökningar kan göras samtidigt i Miljö nätet och i olika databaser runt om i Sverige och i övriga världen. Detta ska lösas genom ett Z39.50-gränssnitt.”<sup>90</sup>

Miljö nätet tillämpar flera klassifikationssystem. Dessa är bl.a. EU CORINE Land Cover för ”Naturtyper” och SCB:s<sup>91</sup> Svensk Näringsgrensindelning för ”Samhällssektorer”, ”Miljö hot” och ”Arter” är också klassifikationssystem. Miljö nätet finns även i en engelsk version och kallas då för Environet.

## 5.2.2 Kvalitetskriterier

Miljö nätet kräver medlemskap för att ta in webbresurser. Det är organisationer, inte privatpersoner, som kan bli medlemmar. Kraven för potentiella medlemsorganisationer är höga och väl formulerade. För att bli beviljad medlemskap ska organisationerna publicera:

- “miljö data, miljö värden eller andra uppgifter om den yttre miljön. Data kan var biologiska, kemiska, fysiska, geografiska, hydrologiska, oceanografiska. meteorologiska, geologiska etc.
- miljö statistik är miljö data som sammanfattas genom olika statistiska metoder. Gränsen mellan data och statistik är flytande.
- miljö information är bearbetade data och statistik om miljön, men också författningar på miljö området, sammanfattningar, bakgrunder till olika miljö frågor, nyheter, utredningar, goda exempel och analyser.”<sup>92</sup>

Dessa ska vara av relevans för miljö arbetet i Sverige. Publiceringen bör innefatta mer än några enstaka dokument av miljö intresse samt beskrivning av mer än bara den egna verksamheten. Myndigheter ska dessutom publicera föreskrifter och annan information med allmän rättsverkan på Miljö nätet. På Svenska miljö nätets webbplats finns ett formulär som ska ifyllas, då en organisation ansöker om att bli medlem. Varje medlemsorganisation ska uppge en kontaktperson, som ska vara insatt i organisationens webbplats, innehåll, administration etc. Medlemskap beviljas av Miljö nätets råd, som gör en bedömning av

<sup>90</sup> <URL: <http://smn.environ.se/miljonat/om-smn/om-smn.htm>> 981230

<sup>91</sup> Se ordlistan.

<sup>92</sup> <URL: [http://smn.environ.se/miljonat/medlinfo/r\\_medlkr.htm](http://smn.environ.se/miljonat/medlinfo/r_medlkr.htm)> 981230

medlemmens webbplats. När medlemskap beviljats bör kontaktpersonen (gärna fler inom organisationen) genomgå en introduktionsutbildning. I denna ingår undervisning hur webbdokument ska anpassas till Miljönätet, klassificering av dokument, samt metadatamärkning. Utbildningen är kostnadsfri.

När medlemskap väl beviljats finns uttalade regler för hur resurserna ska vara uppbyggda. De dokument som Miljönätet länkar till ska uppfylla följande krav:

- ”vara godkända för publicering av organisationens ledning
- ha en länk till organisationens välkomstsida och ha lättillgängliga uppgifter om informationsansvarig, adress, telefonnummer, faxnummer samt e-postadress
- ha minst samma aktualitet som organisationens tryckta information
- ha uppgifter om när informationen senast uppdaterades
- ha länkar och innehåll som regelbundet kontrolleras
- ha innehållsbeskrivningar (metadata) enligt Miljönätets standard och normalt vara försedda med Miljönätets logotype
- inte innehålla reklam, med undantag för sådan som berör medlemmens egen information
- vara utformade så att svarstiderna under normala förhållanden blir rimligt korta
- ha ett tillfredsställande skydd mot att manipuleras av obehöriga.”<sup>93</sup>

När medlemmarna vill bidra med ett dokument till Miljönätet ska kansliet meddelas var detta finns att hämta. Dokumentet ska då var helt anpassat till kraven samt försedd med metadata.

### 5.2.3 Inmatningsformulär

Här följer en översiktlig redogörelse över vilka fält Svenska miljönätet tillämpar i sitt inmatningsformulär:

DC:s fältnamn	Miljönätets fältnamn	Kommentar
Title	Titel	Titel på svenska eller engelska i två separata fält.
Publisher	Medlem	Meny med medlemsorganisationernas namn.
Type	Dokumenttyp	Meny med olika typalternativ.
Language	Språk	Meny med språkalternativ. Svenska och engelska.

<sup>93</sup> <URL: [http://smn.viron.se/miljonat/medlinfo/r\\_kvalit.htm](http://smn.viron.se/miljonat/medlinfo/r_kvalit.htm)> 981230

Subject	Katalogingångar/klassificeringsprinciper	Klassificering med SMN.
Creator	Författare	
Subject: keyword	Fria nyckelord	Nyckelord på svenska eller engelska i två separata fält.
Description	Sammanfattning	Sammanfattning på svenska eller engelska i två separata fält.
Date	Publikationsdatum	
Identifier	Visa - URL	

Följande fält är inte aktivt inbegripna i inmatningsformuläret, men kan användas vid behov. Dessa fält blir dock inte sökbara i databasen, utan ifylles om man önskar få dem sökbara genom andra metadatakänsliga sökmotorer.

DC:s fältnamn	Miljönätets fältnamn	Kommentar
Contributor	Contributor	
Format	Format	
Source	Source	
Relation	Relation	
Coverage	Coverage	
Rights	Rights	

## 5.3 SVESÖK<sup>94</sup>

### 5.3.1 Innehåll och målsättning

Svesök är Kungl.bibliotekets satsning på en rikstäckande söktjänst med Internetresurser. Innehållet är till för allmänheten och inte inriktat på något speciellt ämne. Syftet är att Svesök ska fungera som en nationalbibliografi över den svenska delen av Internet och som ett redskap att katalogisera webbsidor. Innehållet ska ha anknytning till Sverige, någon svensk person, det svenska språket eller finnas på en svensk server. Webbsidorna ska ha adress som slutar på “.se”. Sidor som ligger på servrar med “.com”, “.org”, “.net” och “.edu” och som är registrerade med svensk postadress accepteras också. Projektet startades i september 1997. Idén till Svesök härstammar från en dansk motsvarighet vid namn Webguiden<sup>95</sup>. Den 12 maj 1998 presenterades en demoversion av Svesök på Internet.

Svesök innehåller en söktjänst och en bläddringsfunktion kallad Länkkatalogen. Båda dessa är sökbara samtidigt. Använder man sig av söktjänsten kan man alltså söka på allt i hela Svesök. Sökröboten som hämtar webbresurser från landets alla servrar är densamma som

<sup>94</sup> <URL: <http://www.svesok.kb.se/index.html>> 981230

<sup>95</sup> <URL:<http://www.webguiden.dk>> 981230

används av Kulturarw<sup>396</sup> på KB. Roboten behöver några veckor för att ta ett varv runt den svenska delen av Internet. Den huvudsakliga skillnaden mellan söktjänsten och länkkatalogen är att den senare innehåller webbplatser som är katalogiserade samt utvalda enligt Svesöks urvalskriterier. Använder man sig av söktjänsten återfinns man allt sökroboten hämtat från hela Sverige. Den 20 oktober 1998 innehöll länkkatalogen c:a 800 svenska webbresurser. Hela Svesök innehöll vid samma tidpunkt 2 miljoner resurser. Man räknar med att det tillkommer 100.000-150.000 webbresurser varje dag.

Alla webbresurser med metadata i Svesök presenteras med titel, URL, ämnesord och utgivare. I länkkatalogen finns dessutom en mer utförlig katalogpost beskriven. Denna fungerar som metadata i den bemärkelsen att dokumenten blir sökbara i Svesök. Ska ett dokument bli sökbart utanför Svesök måste det förses med inbäddad metadata.

I Svesöks inmatningsformulär har man under fältet "Kontrollerade ämnesord" tillgång till en länk som leder till det elektroniska ämnesordregistret BUS. BUS står för "Enheten för bibliografisk utveckling och samordning" och är KB:s egen utvecklingsavdelning för ämnesordsindexering. BUS följer de principer som rekommenderas av IFLA<sup>97</sup>. Svesök hänvisar också metadataproducenten till en tryckt version av SAB:s ämnesordlista. Denna har man i dagsläget ingen möjlighet att tillhandahålla elektroniskt.

### 5.3.2 Kvalitetskriterier

Inom den del av Svesök som innehåller material inhämtat av roboten finns inga krav på vare sig kvalitet eller metadata. När det gäller länkkatalogen är dock Svesök noga med att resurserna ska vara kvalitetssäkrade. Med detta menas att Svesöks redaktion i samarbete med ämnesexperter valt ut lämpliga resurser. Ämnesspecialisterna är framförallt bibliotekarierna vid ansvarsbiblioteken (universitets- och högskolebibliotek) samt folkbiblioteken i Stockholm, Göteborg och Malmö.

Vem som vill kan ansöka om att få med sin webbresurs i länkkatalogen. Detta görs i samband med att dokumentet märks med metadata. Metadata är ett av kvalitetskraven för att få ett dokument länkat från Svesöks länkkatalog. Andra krav är att webbsidorna ska vara till för allmänheten och inte av alltför lokal, kommersiell och privat karaktär. Man ska lätt se vem som ansvarar för sidorna och uppdatering bör ske kontinuerligt.

Följande rubriker används av Svesöks redaktion för att göra bedömningen angående införlivande av resurser i länkkatalogen:

- "Helhetsbedömning..."
- Bedömning av innehållet...

---

<sup>96</sup> Se ordlistan.

<sup>97</sup> Se ordlistan.

- Bedömning av formen...
- Bedömning av funktion/system...<sup>98</sup>

### 5.3.3 Inmatningsformulär

Här följer en översiktlig redogörelse över vilka fält Svesök tillämpar i sitt inmatningsformulär:

DC:s fältnamn	Svesöks fältnamn	Kommentar
Identifier	Nätadress: URL	Separat fält för enbart URL.
Identifier	Standardnummer	Listrutor.
Title	Titel	
Creator	Författare/textansvarig	
Contributor	Medarbetare	
Publisher	Utgivare	Utgivarens e-postadress i separat fält.
Subject	Kontrollerade ämnesord	Hänvisning till tryckt version av SAB samt länk till det elektroniska ämnesordregistret BUS.
Subject	Egna ämnesord	
Description	Sammanfattning	
Date	Datum då sidan skapades	
Date	Datum då sidan senast uppdaterades	
Type	Typ av resurs	Listrutor.
Format	Format	Listrutor.
Source	Källa	
Language	Språk	Listrutor.
Relation	Relation	
Coverage	Täckning	
Rights	Rättigheter	

Svesök har också ett antal fält som inte härstammar från DC. De genererar ingen metadata. Fälten är:

Fältnamn	Kommentar
Önskemål om placering	I detta fält ska man fylla i ett ämnesord som finns i Svesöks länkkatalog.

<sup>98</sup> <URL:http://www.svesok.kb.se/info/> 981230

Om sidan uppdateras, hur ofta görs det?	Listrutor.
Datum. Om sidan är temporär, när ska den i så fall tas bort ur Svesök?	
Anmärkning	

Svesöks ämnesredaktörer har ett eget inmatningsformulär som inte är tillgängligt för allmänheten, utan återfinns med hjälp av lösenord. Formuläret liknar allmänhetens, men innehåller dessutom ett extra fält för SAB-klassning. Ämnesredaktörerna letar aktivt lämpliga dokument som ska ingå i länkkatalogen. När de hittat en intressant resurs fyller de i inmatningsformuläret på samma sätt som allmänheten, men det aktuella dokumentets metadatamärkning hamnar inte i HTML-koden (eftersom enbart webbproducenten själv har tillgång till ändringar i den), utan det skapas en separat katalogpost. Denna post fungerar som metadata i Svesök.

## 6. RESULTATREDOVISNING

I detta kapitel diskuterar vi de resultat vi uppnått när vi undersökt projekten med våra frågeställningar som bakgrund. Kapitlet är indelat i tre avsnitt i vilka vi undersöker samordningsförsök, användarvänlighet och kvalitetskontroll. Resultatet är baserat på intervjuer med projektansvariga, en enkät om användarvänlighet samt projektbeskrivningar på Internet. Under varje rubrik redovisar vi i tur och ordning vad som råder inom respektive söktjänst. Under rubriken ”Användarvänlighet” diskuterar vi användare i form av metadataproducenter.

### 6.1 SAMORDNINGSFÖRSÖK

#### 6.1.1 Val av metadataformat

SAFARI valde i samarbete med NetLab att använda sig av DC p.g.a. att formatet ansågs ha goda möjligheter att bli en standard. DC:s femton element täckte dessutom SAFARI:s behov. Inom söktjänsten är man stort sett nöjda med formatet, men har kompletterat med fältet “målgrupp”.

Orsaken till att Miljönätet valde att använda sig av DC-formatet var mer eller mindre en slump. Beslutet togs på ett tidigt stadium, innan DC vunnit mark och blivit en påtänkt standard. Ansvariga inom Miljönätet är stort sett nöjda med DC-formatet.

Orsaken till att Svesök valde att använda DC är följande:

- “DC är utvecklat för att användas för katalogisering av elektroniska dokument på Internet.
- DC är ingen standard men på god väg att bli.
- DC är utvecklat av bibliotekarier, nätverkstekniker och databasspecialister.
- DC används internationellt.
- DC kan användas av webbproducenterna själva.
- DC är en enklare form av katalogisering än ex Marc-formatet. Det är en förutsättning för att vi ska ha en chans att hinna med att katalogisera ett urval av den svenska webben.”<sup>99</sup>

DC var relativt nytt då Svesök valde format, därför hade man heller inga uttalade förväntningar på det. Det anses fungera relativt bra. En nackdel som nämns är att systemet inte är stabilt, d.v.s. fortfarande under utveckling, vilket ibland gör det svårt att arbeta med. Samtidigt inser man att bristen på stabilitet ligger i ämnets natur. Webbresurser är ju inte heller stabila.

---

<sup>99</sup> E-postintervju med Lotta Jansson och Anette Svensson på Svesök, december 1998.

## 6.1.2 Lokala varianter och språkbruk

Definitioner och förklaringar:

- I SAFARI finns s.k. öppnade fält vilka är fält man kan öppna om så önskas, men de är egentligen inte del av det aktiva formuläret. Motsvarande finns i Svenska miljönätet där de icke översatta fälten i slutet på formuläret i praktiken fungerar som öppnade fält, varför vi även refererar till dessa som ”öppnade fält”.
- Standardinställningen är den inställning som gäller i fältet när användaren inte gör något annat val<sup>100</sup>.
- Vi jämför SAFARI:s, Miljönätets och Svesöks formulär med ”Nordic DC metadata creator”<sup>101</sup>. Detta inmatningsformulär återspeglar grundidén med Dublin Core varför vi anser det värdefullt att göra dessa jämförelser.
- Vissa fälts innehåll är så att säga slumpmässiga, d.v.s. den som lägger in metadatan bestämmer själv vad som ska stå utan att gå efter en ämnesordlista eller annan standardiserad lista. Så är fallet i t.ex. Title, Creator, Description, Keywords och Contributor. Dessa fält saknar m.a.o. ett standardiserat innehåll varför en jämförelse dem emellan i samsökningssyfte oftast inte är aktuell. Således kommenterar vi inte alltid dessa element i vår jämförelse nedan. I de fall där två av tre söktjänster har fälten som öppnade fält kommenterar vi heller inte samsökningsmöjligheten eftersom dessa fält ändå inte är sökbara inom respektive söktjänst.

### 1. ”Title”

**SAFARI:s namn:** Titel (obligatorisk)

*Ex:* <META NAME="DC.Title" CONTENT="Botaniska exkursioner">

Kommentar: Vad gäller hjälptexten till ”Titel” skriver SAFARI att ”här skriver du in det namn du tycker passar bäst för sidans innehåll.” Denna formulering synes tvetydig. Den verkar vara mer en uppmaning till att beskriva dokumentet istället för att lägga in den av författaren givna titeln. I ”Nordic DC metadata creator” skriver man att ”Titel” är det namn på resursen som författaren eller utgivaren givit den. Bland de öppnade fälten ligger ett ”Alternativ titel”-fält.

**Svenska miljönätets namn:** Titel, svenska  
Titel, engelska (obligatoriskt att fylla i ett av dem)

*Ex:* <meta name="DC.Title" lang="swe" content="Botaniska exkursioner">  
<meta name="DC.Title" lang="eng" content="Botanic excursions">

Kommentar: Endast ett av ovanstående fält bör användas men båda kan användas. I metataggen skapas en automatisk language-qualifier. ”Swe” eller ”eng” beroende på vilket fält man väljer. Det finns inget fält för alternativ titel.

<sup>100</sup> Från engelskans ”default value”.

<sup>101</sup> SAFARI:s formulär: <URL: <http://safari.hsv.se/metadata/safari.pl?lang=sv>> 990130

Svenska Miljönätets formulär: <URL: <http://smn.environ.se/miljonat/dc-gen/dcgen.htm>> 990130

Svesöks formulär: <URL: <http://www.svesok.kb.se/user/submit/>> 990130

Nordic DC metadata creator: <URL: <http://www.lub.lu.se/cgi-bin/nmdc.pl>> 990130



**Svesöks namn:** Titel (obligatorisk)  
Alternativ titel

Ex: <META NAME="DC.Title" CONTENT="Botaniska exkursioner">  
<META NAME="DC.Title.Alternative" CONTENT="Skönläsnings flora">

Kommentar: I "Alternativ titel" finns inga underavdelningar. Man kan t.ex. inte specificera om det handlar om undertitel eller parallelltitel etc.

## 2. "Creator" (Author or Creator)

**SAFARI:s namn:** Författare/textansvarig (obligatorisk)

Ex: <META NAME="DC.Creator.PersonalName" CONTENT="Andersson, Anna">  
<META NAME="DC.Creator.CorporateName" CONTENT="Botaniska institutionen">

Kommentar: Här ligger två type-qualifiers, "PersonalName" och "CorporateName", genom vilka man kan specificera om författaren/textansvarig är en privatperson eller en organisation. I listrutan för dessa två alternativ visas hur man skriver personnamn (Efternamn, Förnamn) men för organisation finns det ingen motsvarande hänvisning (t.ex. om man skriver institutionsnamnet före universitetsnamnet eller tvärtom). Fältet är upprepningsbart. Här finns även ett fält för författare/textansvarigs e-postadress. Inställningen av type-qualifiern för e-postadressen sker automatiskt beroende på vad man lagt in ovan, d.v.s antingen "PersonalName.Adress" eller "CorporateName.Adress".

**Svenska miljönätets namn:** Författare

Ex: <meta name="DC.Creator.PersonalName" content="Andersson, Anna">

Kommentar: Författarfältet tillåter endast egennamn, vilket gör type-qualifiern "PersonalName" till standardinställning. Vid flera författare, anges huvudförfattaren här och övriga författare i fältet "Contributor" vilket ligger längre ner i formuläret som ett öppet fält. Detta fält är inte upprepningsbart. Åsa Präntare skriver dock att fler namn går att skriva in med komma emellan<sup>102</sup>. Denna information står inte att finna i hjälptexten. Det finns inget fält för författarens adress.

**Svesöks namn:** Författare, textansvarig (obligatorisk)

Ex: <META NAME="DC.Creator.PersonalName" CONTENT="Andersson, Anna">  
<META NAME="DC.Creator.CorporateName" CONTENT="Botaniska institutionen">

---

<sup>102</sup> Präntare. Jämförelse...

Kommentar: Type-qualifiers för personnamn och organisationsnamn finns. Det finns inget fält för författarens adress.

### Övergripande kommentar:

Fälten är likadant utformade såtillvida att de alla har en "PersonalName" type-qualifier. Både SAFARI och Svesök har dessutom "CorporateName" type-qualifiers. Svenska miljönätet saknar detta sista alternativ.

## 3. "Subject" (Subject and Keywords)

### 3a Subject, classification and controlled vocabulary

**SAFARI:s namn:** Ämne: Klassifikation CERIF (obligatorisk)

Ex: `<META NAME="DC.Subject" CONTENT="(SCHEME=CERIF) B290 Systematisk botanik, taxonomi, morfologi, kemotaxonomi. V&auml;xters ysiologi (inte k&auml;rlv&auml;xter).">`

Kommentar: CERIF fungerar här som en scheme-qualifier där den förstnämnda bokstavs- och sifferkombinationen är den primära informationen. Det finns även två öppnade fält; "Ämne: Vokabulär" och "Ämne: Klassifikation".

**Svenska miljönätets namn:** Miljöhot

Samhällssektor  
Geografiskt område  
Kemiskt ämne  
Art  
Miljöarbete  
Miljödata/statistik (obligatoriskt att lägga in en av dessa)

Ex: `<meta name="DC.Subject.Species" scheme="SMN" lang="swe" content="&Ouml;vriga v&auml; &times;ter mm">`

Kommentar: Här ska man välja en katalogingång d.v.s. en av de ovanstående ämnesindelningarna, och klicka sig ner i hierarkin. Det finns dessutom en sökfunktion som visar exakt var i hierarkin man ska söka ämnesordet. Inom dessa ämnesingångar kan man ta sig ner flera steg i hierarkin. Underavdelningarna bildar schemes (scheme="SMN" precis som SAFARI:s kontrollerade vokabulär blir SCHEME=CERIF). Själva katalogingångarna, d.v.s. de åtta ämnesindelningarna, kodas som type-qualifiers i metataggen efter DC.Subject (där innehållet automatiskt översätts till engelska, ex. DC.Subject.Species =Art). Avdelningarna som ligger under de åtta ämnesindelningarna i hierarkin blir däremot kodade som scheme="SMN". En automatisk language-qualifier skapas också i metataggen.

Vad som är anmärkningsvärt är "Geografisk område". Detta är egentligen "Coverage"(täckning). Metataggen för detta fält blir också mycket riktigt "DC.Coverage"(för vidare kommentar se avsnitt 14: "Coverage").

Ex: <meta name="DC.Coverage.Spatial.AreaName" scheme="SMN" lang="swe" content="Sk&aring;nel&auml;n">

**Svesöks namn:** Kontrollerade ämnesord

Ex: <META NAME="DC.Subject" SCHEME="SAB" CONTENT="Ekologi">

Kommentar: Ovanför fältet står ”Ange ett ord eller begrepp per fält. Orden ska hämtas ur SAB:s ämnesordslista” I hjälptexten står det vidare ”Beskriv webbplatsens/sidans innehåll med kontrollerade ämnesord (nyckelord) hämtade från SAB:s ämnesordsregister och från BUS:s lista med godkända ämnesord.” Detta kan bädda för missförstånd för den oinvidige. Det står att man ska fylla i med hjälp av SAB men att registret endast finns tillgängligt i tryckt format. Förhållandet mellan SAB och BUS är inte uppenbart. Oavsett om man tar ett ämnesord från den tryckta versionen av SAB eller tar ett ord ur BUS ämneslista så blir SAB scheme-qualifier.

### Övergripande kommentar:

Frågan är om innehållet i de kontrollerade ämnesordlistorna kan mappas mot varandra? Mappning bygger på att systemen är någorlunda jämförbara, vilket i detta fall är tveksamt. CERIF är ett grunt system med 369 termer, SAB innefattar ca 30 000 ord och SMN är begränsat i sin omfattning, men i sin tur högst specialiserat inom sitt område. Ämnesorden i CERIF:s begränsade vokabulär bör i stor utsträckning kunna mappas mot motsvarande ord i SAB. Att mappa från SAB till CERIF lär däremot innebära problem. SAB:s omfattning förhindrar att man finner motsvarande ord i CERIF. SMN och CERIF lär också bli svåra att mappa mot varandra, då SMN är specialiserat på miljöområdet och inte lär finna motsvarigheter i det mer allmänt inriktade och relativt begränsade CERIF. Att mappa SMN mot SAB bör fungera i viss utsträckning, dock inte helt och hållet. SAB innehåller inte lika många specialiserade miljötermer som CERIF. Det är mycket svårt att mappa SAB mot SMN, p.g.a. SAB:s omfattande och allmänna inriktning och SMN:s specialisering.

Anmärkningsvärt är också att själva scheme-qualifiern ligger som en parentes i ”CONTENT” i metataggen i SAFARI medan den ligger direkt efter elementnamnet i Svesök. Svenska miljö nätets taggar har sina scheme-qualifiers på samma ställe som Svesök, men de har å andra sidan sin egen version där de lägger katalogingången direkt efter ”Subject” (som dessutom översätts till engelska).

### 3b Subject, keywords

**SAFARI:s namn** Ämne: Nyckelord

Ex: <META NAME="DC.Subject" CONTENT="Botanik">

Kommentar: Här handlar det om fritt valda nyckelord där inga qualifiers skapas.

**Svenska Miljönätets namn:** Fria nyckelord, svenska  
Fria nyckelord, engelska

Ex: <meta name="DC.Subject" lang="swe" content="botanik">

`<meta name="DC.Subject" lang="eng" content="botany">`

Kommentar: Orden som matas in i dessa fält läggs direkt i metataggens innehåll, d.v.s inga types eller schemes skapas. Dock skapas en automatisk language-qualifier.

**Svesöks namn:** Eegna ämnesord

Ex: `<META NAME="DC.Subject" CONTENT="botanik">`

Kommentar: Orden som matas in i dessa fält läggs direkt i metataggens innehåll, d.v.s. inga types eller schemes skapas.

#### 4. "Description"

**SAFARI:s namn:** Beskrivning (obligatorisk)

Ex: `<META NAME="DC.Description" CONTENT="En uppsats om Sk&aring;nes flora">`

**Svenska miljönätets namn:** Sammanfattning, svenska  
Sammanfattning, engelska

Ex: `<meta name="DC.Description" lang="swe" content="En uppsats om Sk&aring;nes flora">`  
`<meta name="DC.Description" lang="eng" content="An essay about the flora of Sk&aring;ne">`

Kommentar: Här kan man lägga in beskrivning parallellt på de olika språken.

**Svesöks namn:** Sammanfattning

Ex: `<META NAME="DC.Description" CONTENT="En uppsats om Sk&aring;nes flora">`

#### 5. "Publisher"

**SAFARI:s namn:** Utgivare (obligatorisk)

Ex: `<META NAME="DC.Publisher" CONTENT="Botaniska institutionen, Stockholms universitet">`  
`<META NAME="DC.Publisher.Address" CONTENT="bot.inst@su.se">`

Kommentar: Man ombeds att lägga in en organisation som utgivare, men ingen "CorporateName" type-qualifier används i metataggen. Här ligger även ett fält för utgivarens adress. Ordvalet förvirrar. Först används ordet "utgivare", och i fältet under skrivs

”Förläggarens e-postadress”. Det är tveksamt vad som skiljer dessa två ord åt. Varför inte ”Utgivarens e-postadress”?

**Svenska miljönätets namn:** Medlem (obligatorisk)

Ex: `<meta name="DC.Publisher.CorporateName" scheme="SMN" lang="swe" content="Svenska Botaniska F&ouml;reningen">`

Kommentar: Man kommer till ett medlemsregister när man klickar på ordet ”medlem” där man kan välja den organisation man tillhör. Eftersom enbart medlemsorganisationer och inga privatpersoner kan anmäla resurser ligger här en ”CorporateName” type-qualifier som standardinställning och man kan inte göra något annat val, d.v.s. alternativet PersonalName. Här finns inget fält för medlemsadress.

**Svesöks namn:** Utgivare

Ex: `<META NAME="DC.Publisher" CONTENT="Botaniska institutionen, Stockholms universitet">`  
`<META NAME="DC.Publisher.Address" CONTENT="bot.inst@su.se">`

Kommentar: Här ligger även ett fält för utgivarens adress.

## 6. ”Contributor” (Other Contributor)

**SAFARI:s namn:**

Kommentar: Här finns endast ett öppet fält som heter ”Medförfattare”.

**Svenska miljönätet:**

Kommentar: Här finns endast ett öppet fält som heter ”Contributor” och ligger längst ner i formuläret, bland de fält som inte aktivt inbegrips i formuläret.

**Svesöks namn:** Medarbetare

Ex: `<META NAME="DC.Contributor.PersonalName" CONTENT="Olsson, Ola">`

Kommentar: Här finns både ”PersonalName” och ”CorporateName” som type-qualifiers.

## 7. ”Date”

**SAFARI:s namn:** Datum. (Bäst före datum)

Ex: `<META NAME="DC.Date.Valid" CONTENT="(SCHEME=ISO8601) 1999-03-20 TO 1999-04-01">`

Kommentar: Vad gäller ”Bäst före datum” används här en type-qualifier, ”Valid”. Detta fält är inte upprepningsbart. Det finns även ett öppet fält som heter ”Datum (kopplat till skapandet eller tillgängliggörandet av resursen)”. Här ligger en scheme-qualifier som bestämmer hur datum ska skrivas, vilket är enligt standarden ISO8601. Dessutom skapas en datummarkering automatiskt, nämligen ”MetadataLastModified”. Det finns inget fält kopplat till denna metatagg i inmatningsformuläret.

Ex: `<META NAME="DC.Date.X-MetadataLastModified" CONTENT="(SCHEME=ISO8601) 1999-03-20">`

SAFARI meddelar att datumanvändningen ska ändras under 1999. Hur denna kommer att se ut vet man ännu inte.

**Svenska miljönätets namn:** Publikationsdatum

Ex: `<meta name="DC.Date.Creation" scheme="ISO8601" content="&sup1;999-0&sup3;-&sup2;0">`  
(OBS: Siffrorna blir underliga både i klartext och i metataggen vid det tillfälle då detta exempel skapades. I ovanstående exempel ska det stå 1999-03-20.)

Kommentar: Det står: ”Här anges det datum då dokumentet gjordes tillgängligt i sin nuvarande form.” d.v.s här ligger en ”Creation” type-qualifier som standardinställning. Scheme för datummarkeringen är ISO8601, och det finns inte utrymme för något annat.

**Svesöks namn:** Datum då sidan skapades  
Datum då sidan senast uppdaterades

Ex: `<META NAME="DC.Date.Created" CONTENT="1999-03-20">`  
Ex: `<META NAME="DC.Date.Modified" CONTENT="1999-03-20">`

Kommentar: I det övre alternativet används type-qualifiern ”Created”. Anmärkningsvärt är att Svesök använder ordet ”created” istället för ”creation” vilket används oftare. I det undre alternativet används type-qualifiern ”Modified”.

I övrigt har de två egna datumfält: Om sidan uppdateras, hur ofta görs det.  
Om sidan är temporär, när skall den i så fall tas bort ur Svesök

Kommentar: I det övre alternativet kan man välja från en listruta ”det alternativ som bäst anger hur ofta webbplatsen/sidan uppdateras.”, t.ex. en gång i veckan eller en gång i månaden. Detta fält generar ingen metadata. I det undre alternativet ska man ange ”bäst-före-datum”. Detta är ett eget fält vilket heller inte skapar någon metadata, vilket är lite förvirrande eftersom type-qualifiern ”Valid” verkar applicerbar. (SAFARI använder sig av detta alternativ.)

**Allmän kommentar om Svesök:** Här används inga schemes gällande hur man skriver ut datum.

### Övergripande kommentar:

De tre söktjänsterna har olika varianter på datumtemat: SAFARI har DC.Date.Valid, Svenska miljönätet har DC.Date.Creation och Svesök har DC.Date.Created samt DC.Date.Modified varför en samsökning torde vara svår då de datum som används åsyftar olika sorters tidsangivelser. Om man jämför qualifiers kan man konstatera att Miljönätets DC.Date.Creation och Svesöks DC.Date.Created åsyftar samma typ av tidsangivelse och bör därför kunna fungera vid samsökning. Anmärkningsvärt är dock att datum skrivs på samma sätt i alla tre söktjänsterna. Trots att Svesök inte har en scheme-qualifier ska man enligt hjälptexten skriva datum enligt ISO8601-standarden, varför någon större skillnad söktjänsterna emellan inte förekommer.

Varför Svesök inte har gjort en DC.Date.Valid-metatagg av sitt ”bäst-före”-fält är svårt att förstå. I ”Nordic DC metadata creator” står att type-qualifiers inom detta element inte har blivit fastställda ännu, vilket kanske förklarar den ovanstående mångfalden.

## 8. ”Type” (Resource Type)

**SAFARI:s namn:** Typ (obligatorisk)

Ex: `<META NAME="DC.Type" CONTENT="Text.Thesis.Masters">`

Kommentar: Här finns ett eget typverktyg som fungerar ungefär som CERIF där man kan välja vilken typ av dokument det är man beskriver utifrån en lista. Innehållet omvandlas till engelska i metataggen som i ovanstående exempel där de svenska orden ”Examensarbete, magisteruppsats” blir ”Text.Thesis.Masters” i metataggen.

**Svenska miljönätets namn:** Dokumenttyp (obligatorisk)

Ex: `<meta name="DC.Type" scheme="SMN" lang="swe" content="E&times;amensarbete">`

(OBS: Bokstäverna blir underliga både i klartext och i metataggen vid tillfället när detta exempel skapades. I ovanstående exempel ska det stå ”examensarbete”.)

Kommentar: Här ligger en scheme qualifier som definieras som ”SMN” d.v.s i detta fallet Miljönätets egen lista över dokumenttyper. Här läggs dessutom till en language qualifier. Miljönätets lista är mer utarbetad än den som återfinns i ”Nordic DC metadata creator”.

**Svesöks namn:** Typ av resurs

Ex: `<META NAME="DC.Type" CONTENT="Text.Paper.Academic">`

Kommentar: Alternativen i listrutan är en bearbetad översättning av typlistan som återfinns i ”Nordic DC metadata creator”. Inga qualifiers skapas. Informationen omvandlas till engelska i metataggarna som i ovanstående exempel där ”Uppsats (akademisk)” blir ”Paper.Academic” i metataggen.

### Övergripande kommentar:

De tre ovanstående fälten är liknande utformade men producerar väldigt olika resultat i innehållet. Orden i innehållet är svåra att mappa mot varandra. Ugångspunkten var ett examensarbete. Resultatet blev "Text.Thesis.Masters", "Examensarbete" samt "Text.Paper.Academic". Ord som knappast är identiska i betydelse. Det blir förmodligen komplicerat att mappa alternativen mot varandra. Anmärkningsvärt är dessutom att Svenska miljönätet som har så många andra fält översatta till engelska, här har valt att ha fältinnehållet på svenska, medan SAFARI och Svesök enbart har fältinnehållet på engelska (informationen översätts automatiskt vid inmatningen), ett faktum som mappning i och för sig kan lösa, men varför göra det mer komplicerat än nödvändigt. Ett samarbete kring en gemensam typlista verkar vara ett mer praktiskt och tilltalande alternativ. I "Nordic DC metadata creator" rapporteras att namnen på type-qualifiers inte är fastställda ännu, vilket kanske förklarar ovanstående opraktiska mångfald.

### 9. "Format"

#### SAFARI:s namn:

Kommentar:: Här finns ett öppnat fält som heter "Format".

#### Svenska miljönätets namn:

Kommentar: Här finns ett öppnat fält som heter "Format".

Svesöks namn: Format

Ex: `<META NAME="DC.Format" CONTENT="text/html (.htm.html)">`

Kommentar: Här finns en listruta med valmöjligheter vilka motsvarar listrutan som återfinns i "Nordic DC metadata creator".

### 10. "Identifier"

SAFARI:s namn: URL (obligatoriskt)

Ex: `<META NAME="DC.Identifier" CONTENT="http://www.bot.index.su.se">`

Kommentar: Identifier-elementet kallar SAFARI för "URL" vilket egentligen är en scheme-qualifier i Identifier-elementet enligt "Nordic DC metadata creator". Andra sätt att identifiera det beskrivna dokumentet saknas. Dock skapas ingen "URL"- scheme i metataggen. Fältet är dessutom upprepningsbart, vilket Åsa Präntare ifrågasätter<sup>103</sup>.

---

<sup>103</sup> Ibid.





man en listruta som visar de olika alternativen där dessa skapar motsvarande schemes i metataggen.

## 12. "Language"

**SAFARI:s namn:** Språk

*Ex:* `<META NAME="DC.Language" CONTENT="(SCHEME=ISO639-1) sv">`

Kommentar: Här finns en scheme-qualifier som bestämmer hur språkkoden ska se ut: ISO639-1, d.v.s "svenska" förkortas "sv".

**Svenska miljönätets namn:** Språk (obligatorisk)

*Ex:* `<meta name="DC.Language" scheme="ISO 639-2" content="swe">`

Kommentar: Miljönätet har ISO639-2 som scheme-qualifier i motsats till SAFARI som har ISO639-1. Skillnaden innebär att "svenska" skrivs som "swe" i miljönätet och "sv" i SAFARI. I "Nordic Metadata Creator" är ISO639-1 standardinställning.

**Svesöks namn:** Språk

*Ex:* `<META NAME="DC.Language" SCHEME="ISO639-2" CONTENT="swe">`

Kommentar: Standarinställningen är ISO639-2 d.v.s. samma som Svenska miljönätets.

### Övergripande kommentar:

Svesök och Svenska miljönätet har samma scheme vilket innebär att språkkoderna i metataggen ser likadana ut, d.v.s "swe" för svenska i motsats till SAFARI som har ett scheme som innebär att svenska kodas "sv". Det sistnämnda rekommenderas i "Nordic DC metadata creator". Dessa två beteckningar, "sv" och "swe", bör dock lätt kunna mappas mot varandra.

## 13. "Relation"

**SAFARI:s namn:**

Kommentar: Här finns ett öppnat fält som heter "Beroenden".

**Svenska miljönätets namn:**

Kommentar: Här finns ett öppnat fält som heter "Relation".

**Svesöks namn:** Relation:

Ex: <META NAME="DC.Relation" CONTENT="http://www.bot.inst.lu.se">

Kommentar: Här finns ingen listruta och inga schemes används. Som jämförelse kan man nämna att i "Nordic DC metadata creator" finns en listruta där man kan välja olika schemes för URL, ISBN etc.

#### 14. "Coverage"

**SAFARI:s namn:** Täckning (geografisk täckning)

Ex: <META NAME="DC.Coverage." CONTENT="(SCHEME=TGN) 1000097">

Kommentar: Här finns ett täckningverktyg (ung. som CERIF) vilket visar en kontrollerad lista med namn på länder. Listan bildar ett scheme som kallas "TGN" vilket är en modell där informationen omvandlas till siffror. Det finns även ett öppnat fält som heter "Täckning (Rums- och/eller tidsmässig täckning)".

**Svenska miljönätets namn:** Geografiskt område

Ex: <meta name="DC.Coverage. Spatial.AreaName" scheme="SMN" lang="swe" content="Skåninge län">

Kommentar: "DC.Coverage" skapas i metataggen via katalogingångarna. Här har de sitt eget scheme (scheme="SMN") och en type-qualifier "Spatial.AreaName". I "Nordic DC metadata creator" återfinns två type-qualifiers, "PlaceName" och "PeriodName" där det första "PlaceName" verkar motsvara Svenska miljönätets "Spatial.AreaName". Man kan fråga sig varför Svenska miljönätet inte använder sig av "PlaceName" då innebörden synes vara densamma. Det finns också ett öppnat fält i formuläret som heter "Coverage".

**Svesöks namn:** Täckning

Ex: <META NAME="DC.Coverage" CONTENT="Sverige">

Kommentar: Här finns inga schemes eller types, utan metadata genereras endast som "DC.Coverage".

#### Övergripande kommentar:

Dessa fält torde skapa svårigheter vad beträffar samsökning. Visserligen är innehållet i både Miljönätets och SAFARI:s fält standardiserade men Svesök använder sig av fri text. Dessutom är Miljönätets lista en förteckning över svenska orter och områden medan SAFARI:s täckningsverktyg är en lista över länder. Mappning dem emellan är således inte aktuell, utom kanske just i fallet då man i Miljönätet anger "rikstäckande" och i SAFARI "Sverige" som är den enda gemensamma nämnaren i systemen. I "Nordic DC metadata creator" meddelas det att detta element är under utveckling.

## 15. "Rights". (Rights Management)

### SAFARI:s namn:

Kommentar: Här ligger ett öppet fält som heter "Rättigheter".

### Svenska miljönätet:

Kommentar: Här ligger ett öppet fält som heter "Rights".

**Svesöks namn:** Rättigheter

Ex: `<META NAME="DC.Rights" CONTENT="http://www.copyright.su.se">`

Kommentar: Inga kvalifera eller schemas används.

### Övriga kommentarer om SAFARI:

Egna fält: Målgrupp

Ex: `<META NAME="SAFARI.TargetGroup" CONTENT="Public">`

Kommentar: Den inlagda informationen översätts automatiskt till engelska i metataggen, t.ex. standardinställningen "allmänhet/skola" blir "public". Fältet genererar visserligen metadata, dock inte DC-metadata.

### Övriga kommentarer om Svenska miljönätet:

Svenska miljönätet har inga egna fält. De har däremot automatiska language-qualifiers i titelfälten, Medlem, Dokumenttyp, ämnesingångarna och nyckelorden. En annan sak värd att påpeka är att Miljönätet inte använder versaler i metataggarna, i motsats till de två andra söktjänsterna.

Ex: `<meta name="DC.Language" scheme="ISO 639-2" content="swe">`  
jämfört med:  
`<META NAME="DC.Language" CONTENT="(SCHEME=ISO639-1) sv">`  
SAFARI)  
`<META NAME="DC.Language" SCHEME="ISO639-2" CONTENT="swe">`  
(Svesök)

### Övriga kommentarer om Svesök:

Egna fält: Önskemål om placering  
Anmärkning

Kommentar: I det övre alternativet ska man ange under vilka rubriker i Svesöks länkkatalog man vill lägga sitt dokument. Detta fält genererar ingen metadata. ”Anmärkning” används för eventuella meddelanden till Svesöks redaktion. Fältet genererar ingen metadata. Detta fält anses viktigt av Svesöks redaktion och det har också visat sig att det används frekvent av både allmänheten och ämnesredaktörerna.

Två egna datumfält finns dessutom, se avsnitt 7: ”Date”.

### **6.1.3 Inställning till interoperabilitet**

Inom SAFARI ställer man sig positiv till allt som kan förväntas underlätta informationssökarens behov samt det som ökar möjligheterna att sprida forskningsinformationen inom databaserna. En Z39.50-profil är framtagen och skulle redan idag kunna underlätta parallella körningar t.ex. mellan SAFARI och LIBRIS. SAFARI samarbetar med Miljönätet och tillsammans håller de på att utveckla ett inmatningsformulär som ska kunna användas för resurser som passar i båda databaserna. SAFARI har också inlett diskussioner med Svesök om hur SAFARI-märkta dokument ska kunna bli sökbara i deras databas.

Miljönätet arbetar aktivt för att snarast möjligt realisera samsökningsplaner med andra databaser. Z39.50 är redan i dagsläget installerat på deras server och ett gränssnitt är under framtagande. I dagsläget samarbetar Miljönätet med Svesök, SAFARI, Kulturmät Sverige, Skoldatanätet m.fl. Ett gemensamt inmatningsformulär för Miljönätet och SAFARI är under produktion. Man samarbetar också med ETC/CDS (Europeiska miljöagenturens Topic Centre for Catalogue of Data Sources).

I Svesök tycker man idén med samsökningar i databaser är viktig och hoppas på att kunna genomföra dessa mellan LIBRIS och Svesök under 1999 via en Z39.50-server. Planerna har funnits ända sedan Svesök startades och metadataformuläret samt hjälptexter har utformats för att passa detta syfte. Då fälten utformades studerade man hur andra söktjänster gjort i sina formulär, men tanken på att samsöka i dessa fanns inte då. Ansvariga inom Svesök anser det viktigt att följa DC och inte skapa en egen dialekt av formatet. Man ställer sig positiv till framtida samsökningar i motsvarande databaser i utlandet. För närvarande är Svesök och Kungl.biblioteket med i en grupp svenska statliga webbtjänster, varav vissa använder DC. Man har deltagit i möten med flera av dessa, men ännu finns inget konkret samarbete. Inom Svesök hoppas man på en vidareutveckling inom området under 1999. Man tycker det är viktigt att samarbeta med andra om tillämpningar av DC och utformning av metadataformulär.

### **6.1.4 Sökmöjligheter**

I SAFARI:s sökformulär för enkel sökning kan man utföra fritextsökningar. Man bör inte använda mer än ett sökord och det går bra att trunkera. Det avancerade sökformuläret inleds med två fält, vilka går att kombinera. I fälten finns möjlighet att genom listrutor välja metadata-

elementen Författare (Creator), Titel (Title), Utgivare (Publisher), Abstract (Description), Nyckelord (Subject: Keyword) och Domän/URL (Identifier). Man kan också välja att söka på dessa fält samtidigt. Det går att bygga vidare sökningen genom att ytterligare specificera med metadata-elementen Ämnesklassifikation (CERIF) (Subject), Geografisk täckning (Coverage) och Typ (Type) av dokument med hjälp av en listruta. Dessa valmöjligheter innebär att respektive elements hjälpverktyg öppnas. Verktygen är samma som finns i inmatningsformuläret för metadata-inmatning. Verktygen innebär möjligheter att direkt välja alternativ från listor, och det man väljer läggs sig automatiskt i sökformuläret. Vidare kan man från en listruta välja målgrupp för informationen. Valmöjligheterna är: Allmänhet och skola, Industri och organisationer, Forskare och Högskolestuderande. Man kan även välja att söka på alla fält samtidigt. Slutligen specificerar man språket till svenska eller engelska eller båda samtidigt. Det går bra att använda booleska operatörer<sup>104</sup> men inte att trunkera.

I Svenska miljönätets enkla sökformulär kan man göra fritextsökningar. Här söker man på allt som finns i hela Miljönätet, även det som inte återfinns i miljökatalogen. Dokumenten är här sökbara i fulltext. Man kan trunkera och använda booleska operatörer. Om man skriver flera ord efter varandra läggs ett ”och” automatiskt in mellan orden när sökningen auktoriseras. Citationstecken vid sökning på hel fras kan tillämpas samt parenteser vid längre sökmeningar. Även i det avancerade formuläret erbjuds fritextsökning. Man kan dessutom söka på metadata-element. De första åttafälten motsvarar miljökatalogens och metadataformulärets katalogingångar. Dessa är: Miljöhot, Samhällssektorer, Naturtyper, Geografiska områden, Kemiska ämnen, Arter, Miljöarbete och Miljöstatistik. Varje fält är länkat till en lista med underavdelningar. Alla dessa (förutom Geografiska områden) motsvarar DC:s Subject. Geografiska områden motsvarar Coverage. Man kan ytterligare specificera sökningarna genom att välja: Variabeltyp eller Medium. Dessa ord är inte länkade till någon förklarande text. Övriga metadatabaserade sökelement man kan söka på är: Fria nyckelord (Subject, keywords), Dokumenttyp (Type), Språk (Language), Medlem/utgivare (Publisher) och Författare (Creator).

I Svesöks enkla sökformulär kan man göra fritextsökning samt använda booleska operatörer, trunkera, maskera med frågetecken eller söka på fras med hjälp av citationstecken. Samma typ av sökningar kan man utföra i det avancerade formuläret. Dessutom har man möjligheter att söka på metadata-elementen: Titel/ord i titel (Title), Författare/textansvarig (Creator), Utgivare (Publisher) och Ämnesord (Subject). Under Ämnesord kan man specificera i en listruta varifrån ämnesorden ska tas. Man kan välja mellan SAB ämnesord eller Kategori i Svesök eller båda samtidigt. Ett stort antal booleska operatörer kan användas i Svesöks avancerade sökformulär. Man har möjligheter att utföra sökningar enbart i länkkatalogen. Eftersom innehållet i länkkatalogen skiljer sig en hel del från materialet som är insamlat av sökroboten, rankas sökresultaten på olika sätt. Vid fritextsökning används relevansrankning för att de mest relevanta länkarna ska presenteras först. Dokument med metadata rankas då högre än dokument med träff enbart i fritext. Även metadatamärkta dokument rankas olika. Först kommer ämnessidor i Svesök, sedan dokument i länkkatalogen och sist dokument insamlade av sökroboten.

---

<sup>104</sup> Se ordlistan.

## 6.1.5 Granularitet

I SAFARI vill man att i princip alla sidor inom en sammanhängande sajt bör vara märkta med metadata, eftersom de ofta skiljer sig åt ämnesmässigt, har olika författare eller är av olika typer.

Själva idén, enligt projektansvariga, med Miljönätet är att informationssökaren ska komma direkt från träfflistan till det intressanta dokumentet och inte hamna på en medlemsajt och sedan leta vidare på egen hand. Därför vill man att metadata ska finnas på varje sida inom en sajt.

I Svesök råder uppfattningen att det är bra om de viktigaste sidorna på en webbplats innehåller metadata, men det är inte nödvändigt att det finns på varje sida inom sajten.

## 6.2 ANVÄNDARVÄNLIGHET

### 6.2.1 Formulärens utformning

I SAFARI har man valt ett begränsat antal DC-fält i inmatningsformuläret i syfte att underlätta för användarna. De som är extra ambitiösa har dock möjlighet att använda s.k. öppnade fält, vilka ger tillgång till samtliga DC-fält.

SAFARI har följande stödfunktioner i sitt inmatningsformulär:

- Varje element är länkat till en hjälptext.
- Länk till DC:s hemsida.
- Utförlig förklaring hur DC-metadata fungerar.
- Listrutor på flera ställen.
- I fältet "Subject" startas ett hjälpverktyg (klassifikationssystemet CERIF). Man klickar sig ner i ett ämnesträd och väljer det ämnesord som passar för dokumentet ifråga. Ordet hamnar automatiskt i inmatningsformuläret.
- I fältet "Typ" startas ett verktyg i vilket man väljer vilken dokumenttyp som passar för resursen.
- I fältet "Täckning" startas ett verktyg i vilket man väljer geografisk plats för resursen.

Angående CERIF anser ansvariga inom SAFARI att det begränsade antalet termer är lätthanterligt både för den som söker information och för den som klassificerar. Den begränsade mängden bidrar dock också till att vissa forskare inte hittar lämpliga ämnesord. Det har också visat sig att många inte klassificerar på en tillräckligt djup nivå, vilket motverkar syftet att underlätta för informationssökarna. Det finns planer på att CERIF ska vidareutvecklas och

innefatta väsentligt fler termer än idag<sup>105</sup>. Martin Sparr på SAFARI har redogjort för varför valet föll på CERIF:

“För SAFARI:s del fanns framför allt följande hänsyn att ta vid val av en ämnesklassifikation:

- Användarna ska uppleva det som ett hjälpmedel utöver vanlig sökning. Det innebär att det inte får vara för stort eller ha för många nivåer.
- Många olika grupper ska kunna använda det vid klassificering.
- Forskarna själva bör känna igen sina ämnesområden på en rimlig nivå.
- Det bör helst finnas på flera språk med tanke på Sveriges medlemskap i EU och eftersom Internet är ett internationellt medium.
- Det behöver bara täcka forskningsrelaterade ämnen.
- SAB finns inte tillgängligt elektroniskt och är dessutom mycket svårhanterligt med sina 30 000 ord för icke-bibliotekarier.”<sup>106</sup>

I Miljönätets inmatningsformulär används främst ett begränsat antal DC-fält. De fält som Miljönätet inte aktivt uppmanar medlemmarna att använda finns dock tillgängliga för de medlemmar som önskar en komplett uppsättning. Det är också för medlemmarnas skull Miljönätet bara efterfrågar det som anses vara viktigt för informationssökarens återfinnande av dokumenten. Enligt ansvariga för projektet får det inte bli för komplicerat och betungande att registrera metadata.

När man använder inmatningsformuläret i Miljönätet delas webbfönstret in i två delar. I den vänstra ser man själva formuläret, samtidigt som man i det högra hela tiden kan följa hur metataggarna ser ut i färdigt skick.

Miljönätet har följande stödfunktioner i sitt inmatningsformulär:

- Varje DC-element är länkat till en hjälptext.
- Länk till DC:s hemsida.
- Utförlig förklaring om hur DC-metadata fungerar.
- Klickbara menyer på flera ställen.
- Introduktionsutbildning med bl.a. metadata-undervisning.
- Ämnesordregister med katalogingångar. För att välja katalogingång klickar man på den länkade huvudtexten vilket visar underavdelningarna. Det finns även i vissa fall underavdelningar till dessa. När man bestämt katalogingång hamnar ordet automatiskt i metataggen.

Ansvariga inom Miljönätet har lagt sig vinn om att ämnesindelningen ska vara tydlig och överskådlig. En stor thesaurus skulle motverka detta syfte. Man har istället valt att tillämpa delar av ett antal olika ämnesordregister som ska passa just Miljönätets ändamål.

Svesök tillämpar alla DC-fält. Den som vill metadatamärka ett dokument i Svesök kan välja mellan ett kort och ett långt inmatningsformulär. Ämnesredaktörerna har ett eget

<sup>105</sup> E-postintervju med Martin Sparr på SAFARI, december 1998.

<sup>106</sup> Ibid.



inmatningsformulär som inte är tillgängligt för allmänheten. Detta formulär liknar i stort sett det långa formuläret för allmänheten, men innehåller även ett fält för SAB-klassning.

Svesök har följande stödfunktioner i sitt inmatningsformulär:

- Varje DC-element är länkat till en hjälptext.
- Länk till DC:s hemsida.
- Utförlig förklaring om hur DC-metadata fungerar.
- Lustrator på flera ställen.
- Ämnesordlistan BUS. Man kommer till BUS via hjälptexten. Registret presenteras som en lång lista med ämnesord. Dessa är dock inte klickbara, utan man får ta sig tillbaka till inmatningsformuläret och fylla i ämnesordet manuellt.

Enligt Svesök använder sig inga privatpersoner eller icke-bibliotekarier av SAB. Ämnesredaktörerna gör det kontinuerligt.

### **6.2.2 Dialog med användarna**

SAFARI har kontakt med ansvariga för forskningsorganisationerna. Inom varje organisation ser kontakten mellan ansvariga och forskare ut på olika sätt.

Svenska miljö nätverket värnar om dialogen med medlemmarna. Inom Miljönätet diskuteras framförallt klassificeringsprinciper för att undvika missförstånd och problem.

Svesöks redaktion för en kontinuerlig dialog med webbproducenter som skickar e-post, ringer eller ställer frågor i anmärkningsfältet angående problem med inmatningen. De vill veta hur vissa fält ska användas, vilka sidor som ska katalogiseras och om deras sidor kommer med i länkkatalogen.

### **6.2.3 Hjälpfunktioner och användarutbildning**

Inom SAFARI anser man att användarna behöver mer instruktioner och utbildning. Detta kommer också att erbjudas under 1999, åtminstone i skriftlig form.

Miljönätet uppfattar att medlemsorganisationerna förstår och är nöjda med katalogingångarnas struktur samt att inga större problem föreligger angående metadataproduktionen i allmänhet. Introduktionsutbildningen innebär också grundlig utbildning i ämnet. Kontinuerlig kontakt och diskussion kring klassificeringsprinciper anses dock nödvändig och förekommer också.

Svesöks redaktion anser att hjälpfunktionerna skulle kunna utvecklas ytterligare.

### **6.2.4 Användarnas upplevelser**

För att få reda på hur användarna upplever inmatningsformulären i de respektive söktjänsterna gjorde vi en mindre användarundersökning via e-post. De som deltagit i undersökningen från Svesök är ämnesredaktörerna. Svesöks redaktion skickade ut en förfrågan på sin interna lista för ämnesredaktörer varifrån en handfull svarade. Från Svenska miljönätet är användarna de olika medlemsorganisationerna och de vi intervjuat har blivit slumpmässigt utvalda av Svenska miljönätets redaktion. Påpekas bör att representanter för dessa medlemsorganisationer har vid inträde i Miljönätet fått en kortare utbildning vad gäller hur man lägger in metadata. SAFARI:s användare är oftast någon på en institution e.dyl. där vederbörande är speciellt ansvarig för metadataproduktionen. Namn har vi fått från våra handledare på NetLab. Vi har dock tagit bort några intervjuer för att antalet intervjusvar från varje söktjänst skulle bli lika. Denna utsortering har gjorts slumpvis.

Angående frågornas utformning var hälften flervalfrågor (fråga nummer 1, 2, 6, 7A och 7B). Av dessa var svarsalternativen i fråga 1, 2, 7A och 7B ”mycket bra, ganska bra, varken bra eller dåligt, ganska dåligt och mycket dåligt”. I fråga 6 var svarsalternativen ”mycket tydligt, ganska tydligt, varken tydligt eller otydligt, ganska otydligt och mycket otydligt”.

På flervalfrågan (Nr 1) ”Hur tycker du inmatningsformuläret fungerar?” var svaren ganska varierande men övervägande positiva. Nio av tolv svarade ”ganska bra” eller ”mycket bra”. SAFARI-användarna svarade dock på skalan från ”mycket bra” till ”ganska dåligt”. Det kommenterades att det inte är lätt för en nybörjare att förstå hur formuläret ska användas och att det tar tid att lägga till nya fält, t.ex. keyword-fält. Vederbörande föreslår att man där lägger in tre tomma fält på en gång. Svenska miljönätet och Svesök fick lite bättre betyg. De flesta, tre av fyra, svarade att formuläret fungerade ganska bra. Det kommenterades att det har varit tekniska problem med överföringarna, där vissa fält trillat bort och sammanfattningarna kapats.

På flervalfrågan (Nr 2) ”Hur tycker du hjälpfunktionerna fungerar?” var svaren övervägande positiva. Nio av tolv svarade ”ganska bra” eller ”mycket bra”. Bäst betyg fick SAFARI där alla svarade ”ganska bra” eller ”mycket bra”. Av Svenska miljönätets användare svarade hälften ”varken bra eller dåligt” och andra hälften ”ganska bra” respektive ”mycket bra”. Det kommenterades följande: ”Hjälpfunktioner är alltid svåra. I många fall får man aldrig svar på det man undrar över utan alltid något annat.” Av Svesöks användare svarade hälften att det fungerade ”mycket bra” medan den andra hälften tyckte det fungerade ”ganska bra” respektive ”mycket bra”. Man upplevde det som mycket positivt att kunna diskutera och ha synpunkter på formuläret.

På frågan (Nr 3) ”Tycker du något fattas i inmatningsformuläret, i så fall vad?” svarade åtta av tolv nej på frågan och resten hade varierande kommentarer. SAFARI:s användare efterlyste bl.a. en struktur som skulle göra det lättare för ovana användare att mata in uppgifter, samt en lista över alla koderna som används. Svenska miljönätets användare svarade samtliga nej på frågan. Svesöks användare efterlyste fler hjälpfunktioner, datum för inmatning och ändringar samt ett fält för ämnesredaktörens signatur.

På frågan (Nr 4) ”Tycker du något är överflödigt i inmatningsformuläret, i så fall vad?” svarade nio av tolv nej på frågan och resten hade varierande kommentarer. SAFARI:s användare kommenterade att ”Typ” var marginellt överflödigt. Vederbörande anmärkte även att ordflödet var överflödigt och uttryckte sitt ogillande över det komplicerade ”akademiska” språket. Svenska miljönätets användare tyckte samtliga att inget var överflödigt. Svesöks användare kommenterade att sammanfattningarna kunde bli onödigt långa och att datumfälten oftast förblev tomma, framför allt ”datum då sidan skapades”.

På frågan (Nr 5) ”Finns det fält du sällan eller aldrig använder, i så fall vilka?” svarade tre av tolv nej på frågan och resten hade varierande kommentarer. SAFARI:s användare kommenterade att de inte använde de öppnade fälten samt ”Geografisk bestämning”. Svenska miljönätets användare kommenterade att de inte använder de frivilliga uppgifterna samt ”Samhällssektorer”,

”Miljöarbete” och ”Miljödata/statistik” (de sistnämnda är dock inga fält utan ämnesalternativ i de olika katalogingångarna.) Av Svesöks användare svarade de flesta, tre av fyra, att de aldrig eller sällan använder ”Täckning”, hälften svarade att de sällan eller aldrig använder ”Källa” och ”Rättigheter”. Andra fält som nämndes var de olika datumfälten, framför allt ”Datum då sidan skapades”.

På flervalfrågan (Nr 6) ”Vad tycker du om språkbruket i formuläret?” visade svaren på allmänt ganska nöjda användare, tio av tolv tyckte att formuläret var ”ganska tydligt” eller ”mycket tydligt”. Av SAFARI:s och Svenska miljönätets användare tyckte de flesta, tre av fyra, att formuläret var ”ganska tydligt” eller ”mycket tydligt”. En av Svenska miljönätets användare tyckte däremot att språkbruket var ”ganska otydligt” och kommenterade följande: ”Problemet i vår organisation är att man måste vara insatt i Internetterminologin när man ska göra inmatningar. I vår organisation är allt arbete med sidorna utlagda på olika enheter, vår hjälp till att producera hemsidor har vi FrontPage och där behöver man inte veta vad taggar är för något för att göra en sida.” Svesöks användare gav bästa betyg åt språkbruket då samtliga svarade ”ganska tydligt” eller ”mycket tydligt”.

På frågorna om de olika klassifikationssystemen i respektive söktjänst var svaren ganska varierande. Flervalfrågan (Nr 7A) till SAFARI:s användare var hur det tyckte CERIF-verktyget fungerade, flervalfrågan (Nr 7B) till Svenska miljönätets användare hur de tyckte katalogingångarna fungerade och frågan (Nr 7C) till Svesöks användare, vilken var lite annorlunda än de andra två frågorna, handlade om huruvida de använde sig av SAB eller BUS. SAFARI:s användare var ganska missnöjda med CERIF. Tre av fyra svarade ”ganska dåligt” och en svarade ”mycket dåligt”. Det hierarkiska systemet kritiserades, att behöva ”...tugga sig ner i systemet. Man kan välja fel väg från början och har ingen överblick.”. Kritik riktades också mot att ”...inte finnas med i systemet...” Vederbörande fortsätter: ”...förstår inte hur man skall kunna anpassa ett europeiskt system till svenska förhållanden utan vidare” och uttrycker vidare preferens för ett förenklat UDK-system. Svenska miljönätets användare var inte lika missnöjda utan snarare relativt nöjda med sina katalogingångar. Hälften svarade ”ganska bra” och andra hälften ”varken bra eller dåligt”, och kommenterade att ”det gäller att man är inne och tittar ofta för att inte glömma bort hur det fungerar.” Av Svesöks användare svarade hälften att de använder SAB och andra hälften att de använder både SAB och BUS.

På frågan om övriga kommentarer fanns, gavs det uttryck för diverse åsikter. Två av fyra av SAFARI:s användare kritiserade återigen CERIF-verktyget: att ämnesindelningen är fullständig inom vissa ämnesområden men ofullständig eller icke-existerande inom andra. Andra kommentarer var bl.a. att det hela går för långsamt och att robotens uppdateringar inte alltid fungerat. Svenska miljönätets användare framförde önskemål om en ”gemensam struktur för metadata”, d.v.s. ett formulär som samtidigt skulle registrera sajten i fler söktjänster än enbart Miljönätet. Vederbörande påpekar att man inte är intresserad av att fylla i ytterligare inmatningsformulär om man skulle vilja anmäla sin resurs till någon annan söktjänst. Övriga kommentarer var att problem föreligger att bestämma huvudingång vad gäller katalogingångarna: ”...så att man ej får en massa parallellgångar.” Vederbörande påpekar vidare skillnaden mellan slutanvändaren, d.v.s. den som söker i söktjänsten, och den användare som lägger in metadata: ”Den som använder söktjänsten har ej samma logik som inmataren. Inmataren matar in så pass sällan så logiken kan bli en annan nästa gång man matar in.” Två av fyra av Svesöks användare kritiserade det faktum att det tar lång tid att få fram nya rader och nya fält i formuläret. Det efterlyses också att förhandsgranskningen skulle visa allt som skrivits in i formuläret och att det skulle kunna gå att spara en halvfärdig sida.

## 6.2.5 Webbproducentens insatser vad beträffar inmatning av metadata

Ansvariga inom SAFARI tycker det är en fördel om webbproducenterna själva matar in metadata. Det är omöjligt för några få utbildade katalogisatörer att klassificera den massiva mängd information som skapas på nätet. En nackdel tror man kan vara att utbildade metadataproducenter saknar objektivitet när det gäller klassifikation.

Inom Miljönätet tycker man det är positivt att webbproducenterna själva ansvarar för metadatamärkning, eftersom det skulle vara omöjligt med en central metadataproduktion som i sin tur skulle kräva en stor central organisation. Miljönätet hade inte kunnat komma igång så snabbt om så hade varit fallet. Dessutom skulle det innebära ett onödigt dubbelarbete om flera organisationer satt och gjorde detta arbete för en och samma resurs. Nackdelarna anser man vara att det är svårt att få klassificeringen homogen.

I Svesök anser man att fördelarna består i att webbproducenten själv kan använda metadata som han/hon har skapat till sin resurs. Webbsidan blir sökbar i alla söktjänster med robotinsamlat material som stödjer metadata, inte bara Svesök. Det finns heller ingen möjlighet för biblioteken att hinna katalogisera alla webbsidor, vilket de kanske skulle vilja göra med visst material. Några nackdelar med att webbproducenterna själva ansvarar för metadatamärkning kan man inte se.

## **6.3 KVALITETSKONTROLL**

### **6.3.1 Kontroll av innehåll och metadata**

Kvalitetskontrollen inom SAFARI går till på så sätt att varje forskningsorganisation ansvarar för vad som läggs ut på respektive server. Hur sedan kvalitetskontrollen inom varje organisation går till varierar. Metadata skapas oftast av någon ansvarig inom organisationen och innehållet kontrolleras av någon som ansvarar för materialet som läggs ut på servern. Vid fel kontaktas webbproducenten och felet åtgärdas. Ibland formulerar webbproducenten själv metadata som ska ingå i dokumentet, medan någon annan sköter HTML-kodning och publicering. Vid flera institutioner har man inte tid att lägga ner resurser på omfattande kvalitetskontroll. Ibland registrerar organisationens webmaster samtliga sidor samt kontrollerar innehållet av resurserna. Centralt fungerar hanteringen manuellt och även där granskas många sidor när de registreras. Det handlar dock inte om någon systematisk kvalitetskontroll.

Med en veckas intervall besöker Miljönätets sökrobot medlemmarnas webbplatser och fångar upp registrerade dokument. När ett dokument hämtas utförs en noggrann kontroll och sortering. Kansliet utvärderar hur det är klassificerat, utreder behov av ändrade klassificeringsmöjligheter samt diskuterar detta med medlemmen. Metadata på de första dokumenten från en ny medlemsorganisation kontrolleras mycket noga. Därefter utförs stickprovskontroller. Kvalitetskontrollen Miljönätet utför är mycket resurskrävande, men momentet anses viktigt och prioriteras därför.

Noggrann kvalitetskontroll av metadata tillhörandes resurser anmälda till Svesöks länkkatalog görs. Metadata som skapas till sidor som inte anmäls kan inte kontrolleras. I övriga Svesök finns inga krav på vare sig kvalitet eller metadata.

## 7. ANALYS OCH DISKUSSION

I detta kapitel för vi en analyserande diskussion kring de resultat vi presenterade i föregående kapitel. Vi diskuterar mot bakgrund av den teoretiska genomgången tidigare i uppsatsen.

### 7.1 SAMORDNINGSFÖRSÖK

När söktjänsterna startades fanns inga direkta tankar eller planer på att man en dag skulle kunna samsöka i dem och därför var heller inte intresset så stort vad beträffar hur andra söktjänster valt att tillämpa DC-formatet. Svesök, som är den yngsta av tjänsterna, studerade visserligen hur några andra söktjänster hade utformat sina inmatningsformulär, men syftet var inte att anpassa sin egen tillämpning för eventuella framtida samsökningar. Ansvariga för söktjänsterna verkar ändå ha varit medvetna om att DC förmodades bli en standard och detta var främsta orsaken till valet av metadataformat. Endast Miljönätet valde DC mer eller mindre p.g.a. en slump och detta kan bero på att man inte uppfattade att DC tenderade att bli en standard vid tidpunkten då beslutet om format fattades. Beslut om att starta ett miljönät fattades 1996 och då var DC ännu mycket ungt och outvecklat. Det bekräftas också inom Nordiska metadataprojektet att DC stabiliserades någorlunda först mellan åren 1996 och 1998. Inom Svesök var man väl medveten om fördelarna, förmodligen just p.g.a. att DC vid tidpunkten för starten av projektet nått en högre utvecklingsnivå. Inom Svesök har också planerna på samsökning i LIBRIS och Svesök funnits ända sedan starten. Det har alltså inom SAFARI och Svesök funnits tankar om fördelarna med att tillämpa ett format som var påtänkt som standard. Att välja ett standardformat underlättar för de samordningsförsök som nu håller på att ta form. De tre söktjänsterna förbereder att inom kort kunna samordna sökningar, dels med varandra och med andra databaser. Det är dock ännu ingen realitet och man vet inte riktigt hur resultatet av försöken kommer att se ut. Alla är dock positivt inställda.

Med samsökning i detta sammanhang menas inte nödvändigtvis att de tre söktjänsterna ska integreras med varandra, utan den inställning som råder inom projekten vad beträffar samsökning i databaser överlag. Det gäller också att hålla isär de två begreppen ”anpassning av inmatningsformulär” och ”samsökning i databaser med hjälp av Z39.50”. En anpassning av två eller flera inmatningsformulär behöver nödvändigtvis inte leda till samsökningar i två databaser, utan kan helt enkelt ha att göra med att man anpassar två söktjänsters formulär till varandra för att underlätta inmatningen för användarna. SAFARI och Miljönätet håller på att ta fram ett gemensamt inmatningsformulär. Men en anpassning av två söktjänsters inmatningsformulär kan också ha som mål att underlätta för en applicering av Z39.50, med målet att samsöka i söktjänsterna och på så vis anpassa sätten att tillämpa metadata. Alla tre söktjänsterna har, eller håller på att installera Z39.50. Av detta kan man dra slutsatsen att samsökningsplanerna är långt framskridna. Detta underbyggs också av Cranfields och Hammars påståenden att Z39.50 slagit igenom på bred front och anses revolutionera informationssökningsprocessen, vilket diskuterades i kapitel 4.4. Protokollet tenderar att bli informationssamhällets sökverktyg nummer ett.

Har då söktjänsterna utformat tillämpningen av DC-metadata i enlighet med DC:s rekommendationer? Och överensstämmer denna tillämpning söktjänsterna emellan i så stor utsträckning att samsökningar är möjliga? Vi har detaljstuderat söktjänsternas sätt att tillämpa DC samt försökt analysera i vilken utsträckning mappning elementinnehållen emellan är möjlig. Vår undersökning visar att de flesta elementen inte vållar några svårigheter. Informationen som användarna uppmuntras mata in i metataggarna stämmer någorlunda väl mellan söktjänsterna och möjliggör mappning. Flertalet element behöver (eller kan) inte mappas i detta sammanhang p.g.a. att innehållet är antingen för generellt eller för individuellt utformat. Det handlar om information såsom egna ämnesord eller där en URL eller ISBN ska matas in. De element vars innehåll inte kräver mappning är:

- Title
- Creator
- Subject, keywords
- Description
- Publisher
- Format
- Contributor
- Identifier
- Source
- Relation
- Rights

De element vars innehåll kan mappas utan större problem är:

- Date
- Language (trots olika standarder bör inte mappning vålla några större problem)

De element vars innehåll lär bli svårt att mappa är:

- Subject, classification och controlled vocabulary (p.g.a. systemens olika omfattning och inriktning samt söktjänsternas olika behov)
- Type (p.g.a. listornas olika utformning)
- Coverage (p.g.a. olika standarder och listor samt söktjänsternas olika behov)

Generellt sett vad beträffar inmatningsformulären finns det likheter mellan SAFARI och Svesök medan Svenska miljönätet har en något avvikande variant. En av skillnaderna är att Miljönätet har engelska parallellfält vid sidan av de svenska. Dessa bildar language-qualifiers i metataggarna, för att beteckna både svenska och engelska. Det kan ha sin förklaring i att deras målgrupp, den stora publiken av miljöintresserade både inom och utom Sverige, är av annan karaktär än SAFARI:s och Svesöks vilka båda vänder sig till intressenter inom Sverige. En annan förklaring är att Miljönätet redan i dagsläget samarbetar med Europeiska miljöagenturens Topic Centre for Catalogue of Data Sources. Man är alltså starkt inriktad på internationellt samarbete. Miljönätet finns även i en engelsk version. I SAFARI och Svesök kan man byta språk för att få hela inmatningsformuläret på engelska, men man använder inte de båda språken parallellt.

Om syftet med samsökningar är att göra lyckade sökningar via Z39.50, kan man ställa sig frågan om söktjänsterna också behöver anpassa sina sökformulär till varandra. I sammanhanget har det inte någon betydelse hur gränssnittet är utformat, eftersom själva meningen med Z39.50 är att informationssökaren ska kunna utföra en sökning utan att komma i kontakt med något annat gränssnitt än det han/hon är van vid och förstår. Däremot måste man ha möjlighet att söka på motsvarande element söktjänsterna emellan. Som informationssökare, med utgångspunkt i t.ex. SAFARI, måste man alltså i en sökning via Z39.50 kunna söka efter samma DC-element som de andra databaserna man är kopplad till kan erbjuda. Det är alltså av största vikt att databaser som intresserar sig för samarbete vad beträffar samsökningar även anpassar sökformulären med detta som utgångspunkt.

Följande DC-element är återsökbara i respektive söktjänst:

DC-element	SAFARI	Miljönätet	Svesök
Title	*		*
Creator	*	*	*
Subject	*	*	*
Description	*		
Publisher	*	*	*
Contributor			
Date			
Type	*	*	
Format			
Identifier	*		
Source			
Language	*		
Relation			
Coverage	*	*	
Rights			

I dagsläget återfinns endast tre gemensamma sökfält hos de tre söktjänsterna. Dessa är Creator, Subject och Publisher. Om söktjänsterna fortskrider med planer om samsökningar, bör sökmöjligheterna i större utsträckning anpassas till varandra.

Vad är då lösningen? Ska söktjänsterna erbjuda informationssökaren ett formulär som möjliggör sökning på alla femton DC-elementen? Vid första anblicken kan det vara lätt att dra en sådan slutsats. Man måste dock ha i åtanke att detta sökformulär i realiteten blir stort och komplicerat. Är informationssökaren mogen för valmöjligheterna ett sådant formulär erbjuder? Det finns ingen given lösning på detta. Det uppstår lätt en konflikt. Å ena sidan bör informationssökaren ha en möjlighet att söka på alla DC-element och kombinera efter önskemål, å andra sidan innebär detta ett komplicerat sökformulär, som kan resultera i att informationssökaren i alla fall inte hittar vad han/hon ska.

Vi har vidare studerat hur väl sökformulären och inmatningsformulären inom söktjänsterna stämmer överens. Kan man söka på de metadata-element man har möjlighet att använda i inmatningsformulären? Detta har kanske inte direkt med samordning att göra men kan, menar vi, vara värt att ta upp till diskussion. Förutom SAFARI utnyttjar inte söktjänsterna fullt ut de möjligheter de har att erbjuda informationssökaren, när det gäller att söka på metadata. Om ett

webbdokument är korrekt märkt med metadata, går viss sådan “till spillo” eftersom det inte är möjligt att återsöka alla element som är inlagda. Vi har dragit slutsatsen att ansvariga inom projekten är medvetna om detta och har anpassat inmatningsformulären med tanke på att all metadata (alltså även den som inte kan sökas inom databasen) ska kunna återsökas av andra metadatakänsliga sökrobotar än i den aktuella databasen. Varför man valt bort möjligheterna att söka på alla inmatade metadata-element beror förmodligen på att söktjänsterna valt att göra ett så användarvänligt sökformulär som möjligt. Det ligger onekligen något i detta då, som vi tidigare tagit upp, ett omfattande sökformulär med många valmöjligheter lätt resulterar i att informationssökaren blir förvirrad och “drunknar” i informationsflödet. Man kan dock göra reflektionen att det är synd att inte utnyttja möjligheterna att återsöka fler metadata-element än de man idag har möjlighet att tillgå. Som exempel kan vi ta Svesöks sökformulär där man endast erbjuds att söka på fyra DC-element, trots att kanske alla femton elementen är inmatade.

En annan aspekt när det rör sig om samsökningar databaser emellan, är vilken inställning söktjänsterna har till granularitet. Det handlar alltså om hur söktjänsterna uppmanar sina medlemmar/metadata-inmatare att märka de olika sidorna inom sajter med metadata. Sökresultaten kommer att bli högst varierande i de olika databaserna vid en eventuell samsökning beroende på om metadata finns på alla sidor inom en sajt, eller enbart på huvudsidan. Uppfattningarna skiljer sig något åt mellan tjänsterna, vilket kan förklaras med att de innehåller olika typer av webbdokument och är uppbyggda på olika sätt. I SAFARI, där man tycker att metadata på varje sida inom en sajt är av vikt, finns många sajter där en institution ansvarar för huvudsidan, som förgrenar sig ner inom flera olika ämnen och är skapade av olika forskare. I ett sådant fall är det naturligtvis av vikt att varje sida innehåller metadata som beskriver just den aktuella sidan. Det är också svårt, för att inte säga omöjligt, att ämnesbestämma och beskriva en huvudsida som förgrenar sig ner till många andra typer av dokument. Inom Miljönätet råder samma inställning och även här fungerar sajterna ofta på liknande sätt som i SAFARI, med en medlemsorganisation som ansvarig för huvudsidan som förgrenar sig vidare till sidor som handlar om olika ämnen och med olika författare. Vitsen är att informationssökaren ska komma direkt till det intressanta dokumentet och inte behöva klicka sig fram och kanske gå vilse på vägen. I Svesök har man en lite annorlunda syn på granularitet. Man tycker i och för sig att det är bra med metadata på de viktigaste sidorna inom en sajt, men att det inte på något vis är nödvändigt på varje sida. Detta kan förklaras med att Svesök fungerar på ett annat sätt än SAFARI och Miljönätet. Det finns i Svesök inget synligt mönster för hur sajterna är uppbyggda, alltså går det heller inte att dra slutsatsen att det är viktigt med metadata på varje sida inom en sajt. Om sajtens olika sidor är relativt homogena finns ingen anledning att vara alltför detaljerad och märka varje sida med metadata.

För att underlätta för informationssökaren i en samsökning bör de ansvariga för databaserna i viss mån anpassa synen på granularitet till varandra. Resultatet kan annars bli att informationssökaren får olika svar på sina sökfrågor. I vissa fall kommer han/hon få detaljerade och direkta träffar, i andra fall indirekta med krav på att användaren själv letar sig vidare inom sajterna. Man kan fråga sig om det är rimligt att begära av metadataproducenten att märka varje sida inom en sajt med metadata. Inom söktjänster där professionella katalogisatörer sköter metadata-inmatningen är detta ett mindre problem, men om man satsar på att få allmänheten att skapa metadata till sina dokument i ett större sammanhang verkar tanken ganska orimlig. Bara



det att få människor att själva ansvara för metadatamärkning är ett bekymmer. Något som enligt Lagoze diskuterades redan på DC-2 (se kapitel 4.2.3.).

## 7.2 ANVÄNDARVÄNLIGHET

Vi inleder denna diskussion med frågan vem som egentligen bör sköta metadatamärkningen. I metadatakretsar menar man att det blir ohållbart i ett större sammanhang att professionella katalogisatörer ska ordna metadata för alla dokument som läggs ut på webben. Handhavandet skulle bli dyrt och tidskrävande. Dessutom skulle det resultera i att metadatan inte blir inbäddad och därför inte sökbar för alla metadatakänsliga sökrobotar. Den skulle lagras i slutna system och bara kunna hittas av dessa systems robotar. Det är därför viktigt att man sprider information och kunskap om metadataproduktion på ett enkelt och lättfattligt sätt. Man har många fördelar att vinna på att informera om metadata till allmänheten och få gemene man att börja märka sina dokument själva. Vi har frågat ansvariga inom söktjänsterna vad de ser för positivt i att webbproducenten ska skapa metadata själv. Samtliga nämner det orimliga i att en central organisation skulle ta på sig detta omfattande arbete. Man nämner också att det blir ett onödigt dubbelarbete om flera organisationer skapar separat metadata till webbdokument för sina respektive söktjänster. Det är alltså varken ekonomiskt eller tidsmässigt försvarbart att någon annan än webbproducenten själv ordnar med metadata. Det negativa som nämns är att en utbildad metadataproducent inte alltid lyckas ämnesbestämma sin resurs på ett korrekt sätt. Detta får naturligtvis negativa konsekvenser vid återsökning av dokumenten. Desto större resurser måste läggas ner på att få metadataproducenten att förstå vikten av att korrekt ämnesbestämma dokument, samt att man satsar på utveckling av användarvänliga hjälpverktyg för kontrollerade vokabulärer o.dyl. Som vi tidigare nämnde i kapitel 4.1 påpekar Hakala m.fl. i rapporten från DC-2 att även om webbproducenten själv skapar metadata till sina dokument, utesluter inte detta att professionella katalogisatörer metadatamärker kvalitetsdokument. Man kan säga att dessa två system att skapa metadata kompletterar varandra.

I diskussionen om vem som bör skapa metadata måste man ta hänsyn till hur de olika söktjänsterna är uppbyggda. Beträffande databaser, där någon form av medlemskap krävs, kan det tyckas rimligt att någon inom respektive medlemsorganisation tar ansvaret för metadatamärkning av organisationens samtliga resurser. Så går det ofta till i SAFARI och Miljönätet. Det är också rimligt att som i Svesök, där man satsar på en kvalitetsutvärderad länksamling, ha utbildade ämnesredaktörer som ansvarar för metadatamärkning. I Svesöks länkkatalog görs dock inte metadatan tillgänglig för andra än just dem som söker i Svesök. Det kan tyckas som ett onödigt dubbelarbete att man inom söktjänsterna skapar separata metadataposter som ingen annan kommer åt, samtidigt som kanske webbproducenten själv skapar inbäddad metadata. Detta är något som också representanter för Miljönätet bekräftar. Frågan är också om inte webbproducenten själv är den rätta att skapa metadata, då han/hon ju är den som bäst känner sitt material.

Alla tre söktjänsterna är måna om att få användarna att förstå hur de ska gå tillväga med inmatningsformulären. De har ansträngt sig för att göra det så lättanvänt som möjligt. Annica

Moshiri på Svenska miljönätet uttrycker ambitionerna vad beträffar utformning av inmatningsformuläret på följande sätt:

”...det ska vara så enkelt, smidigt och snabbt som möjligt för våra medlemmar att registrera sin metadata.”<sup>107</sup>

Tanken med DC är att användarna själva ska kunna beskriva sina dokument och de tre projekten har försökt uppnå detta syfte var och en på sitt sätt. För att få webbproducenter att använda sig av inmatningsformulären krävs att de framställs på ett pedagogiskt och relativt enkelt sätt. I denna diskussion får man ha i åtanke att den bild som vår undersökning ger av de tre söktjänsterna inte riktigt motsvarar hur DC förhoppningsvis kommer att se ut i ett större sammanhang i framtiden. Vi tycker dock att man kan se på denna undersökning som ett försök att i detta relativt begränsade sammanhang se vad som kan vara av vikt att tänka på när man skapar ett inmatningsformulär. Dessa tidiga försök att skapa användarvänliga inmatningsformulär kan komma till nytta när man senare försöker introducera DC på bred front. Då gäller det att kunna tillhandahålla pedagogiska formulär så att användarna inte skräms bort utan tar konceptet till sig.

Frågan är då hur väl söktjänsterna har lyckats med att skapa användarvänliga inmatningsformulär. Vi har studerat söktjänsternas formulär, vilka stödfunktioner som tillhandahålls samt genom intervjuer med ansvariga tagit reda på om någon dialog med användarna existerar, samt vilka brister som eventuellt finns i samband med metadatomärkning. Detta har vi sedan ställt mot resultaten från användarundersökningen. Man bör ha i åtanke att SAFARI och Miljönätet har anpassat sina inmatningsformulär till sina respektive ämnesområden, medan Svesöks är av mer allmän karaktär.

Faktum är att alla DC-fält tillämpas i de tre inmatningsformulären. SAFARI och Miljönätet har dock valt att aktivt inbegripa ett begränsat antal, medan de fält som anses mindre viktiga är placerade längst ner i formulären. Dessa fält är inte heller sökbara genom respektive söktjänsts sökformulär, utan kan bara återfinnas genom andra metadatakänsliga sökrobotar. Fälten används också relativt sällan av användarna. Svesök erbjuder användarna möjligheten att fylla i antingen ett långt formulär med alla fält inbegripna, eller ett kort formulär med ett begränsat antal metadatafält. Syftet med detta är att underlätta för användarna. Ett fullständigt inmatningsformulär kan, liksom ett fullständigt sökformulär, snarare avskräcka än underlätta för en ovan användare. Ämnesredaktörerna, vilka vi intervjuat, räknas dock inte till ovan användare och fyller alltid i ett långt formulär.

Av användarna inom söktjänsterna verkar de flesta relativt nöjda med hur formulären fungerar. Speciellt Miljönätets användare är nöjda, möjligtvis med någon invändning mot språkbruket. De flesta är nöjda med antalet fält som erbjuds i formulären. I SAFARI reagerade någon på att fältet Typ var lite överflödigt. I Svesök reagerade flera på att så många datumfält tillämpades. Täckning, Källa och Rättigheter nämns också av flera som fält vilka sällan används. I SAFARI och Svesök är användarna nöjda med språkbruket.

---

<sup>107</sup> E-postintervju med Annica Moshiri på Svenska miljönätet, december 1998.

Varje söktjänst har ett antal stödfunktioner som är till för att underlätta för användarna. Dessa liknar varandra till stor del i form av hjälptexter, menyer o.dyl. SAFARI och Svesök använder genomgående listrutor, medan Miljönätet har valt att länka till nya sidor med menyer på flera ställen. SAFARI erbjuder hjälpverktyg på tre ställen: ämnesverktyget CERIF, ett typverktyg och ett täckningsverktyg. Svesök länkar till BUS ämnesordregister. Av ämnesredaktörerna förväntar man sig att de ska lägga in ämnesord från den tryckta versionen av SAB. I SAFARI är användarna överlag nöjda med stödfunktionerna, men däremot återfinns en hel del klagomål på CERIF. Forskarna hittar inte lämpliga ämnesord i registret. Enligt Martin Sparr på SAFARI ska CERIF utvecklas vidare och detta välkomnas säkert varmt av användarna. Miljönätets användare är mer nöjda med möjligheterna till klassificering genom katalogingångarna, samt är relativt nöjda med övriga stödfunktioner. Svesöks ämnesredaktörer efterlyste fler stödfunktioner, men verkar ändå relativt nöjda. Man kan se ett visst mönster som bekräftar att användarna är mer nöjda än missnöjda med utformningen av inmatningsformulären. Av detta drar vi slutsatsen att söktjänsterna har lyckats relativt väl med att göra formulären användarvänliga.

Alla söktjänsterna för någon typ av dialog med användarna. Kommunikationen ser dock lite olika ut, beroende på att söktjänsterna helt enkelt är uppbyggda på olika sätt. Inom SAFARI sker kontakten i flera steg, eftersom varje forskningsorganisation har någon som ansvarar för att metadata läggs in korrekt och som också ansvarar för kvaliteten på dokumenten man bidrar med. Tyvärr verkar det finnas vissa brister i denna flerstegskommunikation. Vi har t.ex. hittat exempel på missförstånd inom SAFARI. Frågan är hur pass insatta de ansvariga är när det gäller metadatamärkning. De har många andra arbetsuppgifter och anser sig ibland inte ha tid att fördjupa sig i handhavandet med metadatamärkning. Följande citat från en ansvarig webmaster på en forskningsinstitution speglar problemet:

”Eftersom SAFARI-projektet är tänkt att hinnas med ovanpå mina andra arbetsuppgifter ... finns det helt enkelt inte tid för någon omfattande kvalitetskontroll.”<sup>108</sup>

I Miljönätet förs dialogen med användarna i viss mån på liknande sätt. Redaktionen står i förbindelse med ansvariga inom medlemsorganisationerna. P.g.a. noggrann kvalitetskontroll minimeras dock risken för att missförstånd slipper igenom. Metadataproducenterna kontaktas och man diskuterar problem innan materialet läggs ut på webben. När det gäller Svesöks ämnesredaktörer är de mycket nöjda med kommunikationen med Svesök. Allmänna webbproducenter kontaktar också ofta redaktionen med frågor och detta är något Svesök uppmuntrar till. Slutsatserna man kan dra av detta är att en kontinuerlig och levande dialog med användarna samt kontroll av resurser borgar för ett gott resultat vad beträffar användarnas metadataproduktion. Frågan är om denna kontakt och kontroll är rimlig och genomförbar i ett större sammanhang.

Ansvariga inom söktjänsterna är medvetna om att det föreligger vissa svårigheter för metadataproducenterna att använda inmatningsformulären. Inom SAFARI satsar man på förbättringar vad beträffar instruktioner och utbildning. Inom Svesök råder också uppfattningen att hjälpfunktionerna skulle kunna utvecklas ytterligare. Ansvariga inom Miljönätet, vilka har

---

<sup>108</sup> E-postintervju med tre forskningsinstitutioner kopplade till SAFARI, februari 1999.

satsat mycket på utbildning, upplever att det inte råder några större problem när det gäller metadatamärkning från medlemsorganisationerna. Utbildning av metadataproducenter kan alltså tyckas vara en bra satsning, men frågan är även här om det är ekonomiskt och tidsmässigt försvarbart i ett större sammanhang.

Hur bör man då gå tillväga för att skapa ett användarvänligt inmatningsformulär? Ett formulär som verkligen får användarna att ta sig an uppgiften att metadatamärka sina dokument. Det gäller ju inte enbart att påtala fördelarna webbproducenten vinner med att skapa metadata till sin resurs, utan även att få honom/henne att förstå hur man ska göra. Formuläret ska vara så lättfattligt att användarna inte behöver gå igenom någon regelrätt utbildning. Processen får inte ta för mycket av den ovana metadataproducentens tid. Det får alltså inte vara för omfattande. Samtidigt får det inte bli så förenklat att metadatan inte fungerar vid återsökning. Det är som synes inte någon lätt uppgift. Här drar vi vissa paralleller med minimalist-strukturalistdiskussionen som pågick under DC-4. Minimalisterna som förespråkade ett enkelt, lättfattligt sätt att tillämpa DC, med huvudsyftet att allmänheten ska kunna ta det till sig, dock med risk för att metadatan inte blir av så hög kvalitet. Strukturalisterna ville istället satsa på en omfattande, detaljerad tillämpning av DC med fördelen att metadatan blir högkvalitativ, men med resultatet att förmodligen bara professionella katalogisatörer skulle till fullo kunna förstå hur inmatningen skulle ske. En medelväg mellan dessa synsätt vore förmodligen den bästa lösningen. Metadata får inte gå förlorad i ansträngningen att uppnå användarvänlighet och användarvänligheten får inte gå förlorad i ansträngningen att få fram högkvalitativ metadata.

### 7.3 KVALITETSKONTROLL

Hur viktig är kvalitetskontrollen i sammanhanget? Måste webbdokumentens innehåll vara av hög kvalitet och vem ska i så fall avgöra vad som är hög kvalitet? Idén med Internet är att vem som helst ska kunna göra sin resurs synlig och naturligtvis kan man inte censurera allt man anser värdelöst. Först och främst måste man dela upp diskussionen om kvalitetskontroll i två underavdelningar: kontroll av kvaliteten på själva resursen samt kontroll av kvaliteten på inlagd metadata. Resursens innehåll anser vi vara av mindre vikt, beroende på vilka kvalitetskriterier som råder inom söktjänsterna. Det är sedan upp till varje söktjänst att se till att kvalitetskontrollen motsvarar kvalitetskriterierna. Vi koncentrerar vår diskussion till kvalitetskontroll av metadata.

Vilka inställningar har man till kvalitetskontroll inom söktjänsterna? Hur stora resurser har man möjlighet att satsa? Inom SAFARI ligger ansvaret för kontroll av kvalitet, både vad beträffar resursernas innehåll samt metadatamärkning på ansvariga inom varje forskningsorganisation. Hur kontrollen av inlagd metadata i realiteten fungerar på respektive institution varierar och den ger intryck av att inte vara alltigenom välstrukturerad och väl fungerande. Brist på tid och engagemang resulterar tyvärr i en inte alltid korrekt metadatamärkning. Vid sökningar i SAFARI stöter man ibland på missförstånd och icke korrekt inlagd metadata. Som exempel på de mildare missförstånden kan nämnas följande: På webbsidan <<http://www.socforsk.se/projekt/aldre.htm>> (99-03-02) ligger Socialmedicinska forskningsrådet som "Personal name" i metataggen. Man har alltså använt sig av "Person som upphovsman" i stället för "Organisation som upphovsman". På

webbsidan <<http://info.adm.umu.se/SAFARI/Projekt/42.shtml>> (99-03-02) har man lagt in Umeå universitet i en Subject-tag. Detta resulterar i att materialet inte blir effektivt sökbar. Viss kontroll sker också centralt inom SAFARI, men denna är inte systematisk på något vis. Miljönätet lägger ner mycket resurser på kvalitetskontroll. Många resurser granskas och sorteras. Vid fel tar kansliet kontakt med ansvariga och missförstånd reds ut. Man kontrollerar metadata från nya medlemmar extra noga. Bland etablerade medlemmar utförs bara stickprovskontroller. Miljönätet medger att detta moment är mycket resurskrävande, men resultatet blir högkvalitativ metadata. I Svesöks länkkatalog kontrolleras naturligtvis det intellektuella innehållet mycket noga. Det är själva villkoret för att resurser överhuvudtaget ska få vara med. Det mesta materialet är metadata märkt av utbildade ämnesredaktörer. Deras metadata märkning kräver ingen kontroll. Övrigt material som plockas upp av sökroboten har varken krav på innehållsmässig kvalitet eller metadata.

Missförstånd och missbruk kring metadata märkning omkullkastar hela idén med att skapa ett effektivt sätt att strukturera och återvinna information. Om irrationella ämnesord läggs in i metataggen, för att webbproducenten vill nå ut till fler kategorier av informationssökare, motverkas syftet med fältsökningar för att få logiska träffar. Samma eller liknande brusnivå som utan metadata kommer att bli konsekvensen. Vi har t.ex. hittat resurser (ej inom SAFARI, Miljönätet och Svesök) där webbproducenten lagt in ett femtiotal nyckelord som inte alls stämt överens med resursens innehåll. Nyckelorden var bl.a. Sex, Pamela Anderson och Mat. Ord som ofta används i sökningar på webben, vilka uppenbarligen är inlagda för att locka fler informationssökare att besöka sidan. Enligt Hakala, Husby och Koch är tillämpning av metadata det viktigaste steget för att utöka kvaliteten och precisionen när det gäller automatisk resursåtervinning. Detta är en sanning med modifikation. Metadata fungerar utmärkt i en kvalitetskontrollerad miljö (kontroll av metadata). Detta innebär visserligen ett kostsamt och tidskrävande arbete, men utan någon form av kontroll tvivlar vi på att konceptet kommer att fungera. Om metadata slår igenom, om man lyckas skapa en standard som är någorlunda enhetlig och om man lyckas få gemene man att inse värdet av metadata, kan man spekulera vilken effekt metadata får på webben i framtiden. I kapitlet om Dublin Core i framtiden diskuterar vi kring Thieles spekulation om inte metadata kommer att dela upp webben i två delar, en akademiskt inriktad och en fri sådan, där den akademiskt inriktade delen använder sig av metadata medan den andra inte gör det. Vidare spekulerar han kring det faktum att användningen av metadata i sig blir ett kvalitetsmärke. Vi har dragit liknande slutsatser. Det faktum att metadata fungerar bäst i ett kvalitetskontrollerat sammanhang bäddar naturligtvis för att metadata kommer att stå för någon sorts kvalitetssäkring.

## 8. SLUTSATSER

Vi tror att olikheterna mellan söktjänsterna vad beträffar DC-metadata kommer att utjämnas med tiden, eftersom det verkar ligga i de ansvarigas intressen att så uniformt som möjligt utforma formulär och metatagging. I detta ligger dock en konflikt som härrör ur själva grundidéen om Dublin Core. De tre undersökta söktjänsterna har olika användare, olika ämnesinriktningar och därmed också olika behov vad beträffar metadata. Detta gör att inmatningsformulär, sökformulär, inställning till granularitet m.m. skiljer sig åt, vilket resulterar i stora olikheter mellan söktjänsterna. Flexibiliteten, alltså möjligheterna att använda olika varianter på temat, slår man fast i DC:s målsättning. Olikheterna kan dock resultera i problem både när det gäller användarvänlighet och interoperabilitet. Om utseendet på de olika inmatningsformulären blir alltför individuellt skapas ingen kontinuitet och därmed förlorar man mycket av användarvänligheten. Det ligger i detta begrepp att användaren i viss mån behöver känna igen sig, att han/hon inte ska behöva ta ställning till de idéer och grepp varje inmatningsformulär uppvisar. Därav behovet av en standard utifrån detta perspektiv.

Vad beträffar samsökningar står det också i DC:s målsättning att formatet ska vara interoperabelt. Detta kanske främst mellan DC och annan metadata men resonemanget gäller även inom det egna formatet, nämligen att en lokal variant av DC-metadata ska fungera ihop med en annan lokal variant. Dessa varianter får inte uppvisa för stora olikheter. Om samsökningar ska kunna realiseras måste innehållet i söktjänsternas olika utformade metataggar kunna mappas mot varandra. Det gäller alltså att uppmuntra användarna att lägga in någorlunda homogen information. Detta mål uppnås genom en relativt homogen tillämpning av metadata och diverse hjälpverktyg.

Vad gäller användarvänligheten anser vi det vara av vikt att söktjänsterna anpassar sina inmatningsformulär så de ger ett någorlunda enhetligt intryck. Även vad beträffar granularitet bör man försöka enas om en slags gemensam standard, även om det ligger i sakens natur att det föreligger skillnader i söktjänsternas sätt att hantera dessa frågor p.g.a. deras inbördes ofrånkomliga olikheter.

Man kan vidare fråga sig om det är rimligt att en användare ska behöva genomgå utbildning för att klara av att fylla i ett inmatningsformulär. I DC:s grundidé finns tanken att formatet ska vara så lättbegripligt att någon utbildning i ämnet egentligen inte bör vara nödvändig alls. Man kan fråga sig om detta är realistiskt. I dagsläget är DC på sina ställen så komplicerad att t.o.m. utbildade katalogisatörer har svårigheter att använda formatet. Därför är det av yttersta vikt att inmatningsformulären utformas så användarvänligt som möjligt.

Man skulle kunna tänka sig ett gemensamt centralt inmatningsformulär som resulterar i att den aktuella webbresursen anmäls till fler än en söktjänst. Tanken på ett sådant formulär ger visserligen upphov till åtskilliga frågeställningar bl.a. hur det skulle se ut. Vem ska t.ex. ansvara för utformningen, minimalister eller strukturalister o.s.v. Man kan också fråga sig hur söktjänster med olika ämnesinriktning ska kunna använda ett och samma formulär. Kan ett gemensamt formulär verkligen täcka allas behov? Vissa söktjänster har dock insett vikten av att driva igenom denna idé. Till dessa hör SAFARI och Svenska miljö nätet som för närvarande håller på att arbeta

fram ett gemensamt formulär för sina användare. Idén med ett gemensamt formulär har sina fördelar. Användarna skulle slippa den ganska krävande processen att lägga in metadata för varje söktjänst de skulle vilja ingå i. I användarintervjuerna uttrycktes önskemål om en gemensam metadatastruktur.

Vad beträffar användarvänligt utformade sökformulär tror vi att korta, begränsade formulär är att föredra framför långa, komplicerade. De senare tenderar att skrämja bort ovana användare. Vi efterlyser dock flera varianter av sökformulär. Idag finns det hos praktiskt taget alla söktjänster ett formulär för enkel sökning och ett för avancerad. Det enkla sökformuläret erbjuder fritextsökning och det avancerade brukar innehålla ett antal fält mellan vilka man kan applicera booleska operatörer. Dessa formulär tjänar visserligen sitt syfte, men de erbjuder inte möjligheter att söka på alla DC-element. I takt med att dokument med fullständig metadata i allt större utsträckning läggs ut på Internet ligger det också i sakens natur att man borde kunna söka på en komplett uppsättning metadata-element. Varför inte skapa ett tredje sökformulär? Förutom enkel och avancerad sökning skulle man även kunna erbjuda informationssökaren att göra en DC-sökning. I detta högst avancerade sökformulär borde alla femton elementen vara sökbara. Varför annars ha avancerad metadata på dokument om man ändå inte kan söka på alla element?

Vad gäller kvalitetskontroll av metadata föreligger en konflikt. Det verkar som om valet står mellan ”dyr” och ”billig” metadata. Dyr blir den när man satsar på noggrann kvalitetskontroll och utförlig användarutbildning. Billig när man inte alls kvalitetskontrollerar, utan låter användarna skapa metadata bäst de själva kan. I det första fallet skapas visserligen högkvalitativ metadata, men förmodligen krävs då professionella katalogisatörer. Resultatet av detta är att kostnaderna blir höga och frågan är om idén överhuvudtaget är genomförbar. Att satsa på det billiga alternativet innebär att man måste acceptera sämre kvalitet där risken för missförstånd och missbruk föreligger. En kompromiss mellan dessa ytterligheter torde vara den väg man bör ta för att göra metadata meningsfull men ändå rimlig.

Vi tror i likhet med Thiele att om DC verkligen slår igenom kommer webben att delas upp i två delar, en mer akademiskt inriktad och en fri. På den akademiskt inriktade delen kommer troligtvis metadata bli ett kvalitetsmärke i sig. Vi har dragit slutsatsen att metadata passar bäst i kvalitetskontrollerade miljöer. Idag när webbproducenter till varje pris försöker komma högst upp på söktjänsternas träfflistor, hade metadata antagligen (och har redan) använts som ett redskap i samma dåliga avsiktens tjänst och bruset i sökningarna förblivit konstant. Därför anser vi att metadata inte fungerar på en sådan fri del av Internet. Metadata blir effektivt först då människor är mer intresserade av att hitta och tillgängliggöra relevanta kvalitetsdokument och mindre intresserade av att hamna överst på söktjänsternas träfflistor.

## ORDLISTA OCH FÖRKORTNINGAR

**Adresserbara objekt** - Objekt som är möjliga att identifiera och återfinna på webben.

**Booleska operatörer** - Bl.a. AND, OR, NOT och NEAR. Med hjälp av dessa kan man formulera sökvillkor för sökmotorer. Möjligheter att använda operatorerna ges oftast i söktjänsternas avancerade sökformulär.

**BIBLINK** - Linking Publishers and National Bibliographic Services, ett projekt inom bl.a. UKOLN\*, finansierat av Telematics Application Programme of the European Union Fourth Framework Programme.

**BIBLINK Core** - BIBLINK:s egen DC-variant.

**BUS** - Enheten för bibliografisk utveckling och samordning. Den enhet på KB\* som ansvarar för nya ämnesord. Ett komplement till SAB\*.

**CERIF** - Common European Research Information Format. Det klassifikationssystem SAFARI använder sig av.

**CNI** - Coalition for Networked Information.

**DC** - Dublin Core.

**DESIRE** - Development of a European Service for Information on Research and Education, ett projekt inom UKOLN\*.

**Dokument** - Webbsida\*.

**DSTC** - The Distributed Systems Technology Centre.

**GIF** - Graphic Interchange Format. Bildformat.

**Granularitet** - Finfördelning. I detta sammanhang: hur "djupt" man väljer att metadatamärka en sajt. Man kan t.ex. mata in metadatan enbart på huvudsidan eller gå djupare ner i sajtens förgreningar och metadatamärka varje sida.

**Gränssnitt** (användar-) - Ett programs synliga beteende mot användaren. Det som händer mellan bildskärm och människa.

**<HEAD>** och **</HEAD>** - De HTML\*-taggar som anger dokumentets "huvud". Mellan dessa två ska metadatan klistras in.

**HTML** - Hyper Text Markup Language. Ett markeringsspråk som används för att programmera hypertext, främst för webbsidor\*.

**IFLA** - International Federation of Library Associations and Institutions.

**Inbäddad metadata** - Metadata som är inklistrad i ett webbdokuments HTML\*-kod.

**Inmatningsformulär** - Elektroniska formulär på Webben\* med vars hjälp man skapar metadata till webbdokument. Den inmatade informationen omvandlas till metataggar\* som sedan klistras in i dokumentet.



**Internet** (nätet) - Det globala datornät av tcp/ip-nätverk som omfattar hela världen.

**Interoperabilitet** - Samsökning i olika databaser.

**ISO** - Information Standards Organization.

**KB** - Kungl. biblioteket.

**Klient** - En dator eller terminal som förbinds med en server\* via ett nätverk.

**Kulturarw3** - KB:s sökrobot som söker och bevarar dokument.

**Kvalitetssöktjänst** - En söktjänst vars innehåll är kvalitetsutvärderat.

**Länksamling** - En uppsättning länkar (ofta ämnesindelade) som pekar mot webbresurser.

**Mappa** - Logiska kopplingar mellan element i en mängd och element i en annan mängd.

**MARC** - Machine Readable Cataloguing. Standardiserat katalogiseringsformat som finns i flera nationella varianter t.ex. USMARC och UKMARC.

**Metadatagenerator** - Den funktion som omvandlar metadata till metataggar i HTML\*-kod.

**Metatagg** - Tecken som används för märkning av metadata.

**NCSA** - The National Center for Supercomputing Applications.

**NetLab (LUB)** - Lunds universitetsbiblioteks utvecklingsavdelning som arbetar med utvecklingsprojekt inom digitala bibliotek och nätbaserad informationsförsörjning.

**NLA** - The National Library of Australia.

**Nordic Metadata Project** - Nordiska metadataprojektet. Ett av de största skandinaviska DC-projekten. Knutet till Helsingfors.

**NORDINFO** - Nordiska samarbetsorganet för vetenskaplig information.

**NWI** - Nordiskt Web Index.

**Nätverksprotokoll** - En regelsamling för dataöverföring i nätverksmiljö.

**OCLC** - Online Computer Library Center (Office of Research).

**Pekare** - När data inte används direkt. Pekaren talar om var den ligger.

**PICS** - Platform for Internet Content Selection. Ett system för etikettering av webbsidor. Man kan med hjälp av PICS göra oönskat material på webben otillgängligt.

**PICT** - Grafikformat använt av Macintosh Operating System.

**Protokoll** - Regelsamling. Se även Nätverksprotokoll.

**RDF** - Resource Description Format. En infrastruktur för metadata som är "maskinbegriplig".

**Resurs** - Webbsida\*

**SAB** - Sveriges Allmänna Biblioteksförening. I detta sammanhang det klassifikationssystem som

tillämpas på de flesta svenska bibliotek.

**SBIG** - Subject Based Information Gateway. En ämnesorienterad, kvalitetsutvärderad länksamling.

**SCB** - Statistiska Centralbyrån.

**Semantik** - Läran om betydelsen hos språkliga uttryck.

**Separat katalogpost** - Istället för att klistra in metadata direkt i ett webbdokuments HTML-kod (inbäddad metadata\*), skapas en separat katalogpost med metadata. Från denna finns en pekare\* till det aktuella dokumentet.

**Server** - Central serviceenhet i nätverksmiljö som bearbetar uppdrag från klienten\*.

**SGML** - Standard Generalized Markup language. En internationell standard för textbeskrivning.

**Subelement** - Ett annat namn för qualificern "Type".

**Syntax** - Läran om hur ord förbinds med varandra till högre enheter, d.v.s. ordgrupper, satser och meningar.

**Sökrobot** - Programvara som rör sig på WWW och samlar in resurser. Kallas också webcrawler eller webspider.

**Taggar** - Tecken som används för märkning av data. Ex. på HTML-tag: <TITLE>

**TIFF** - Tagged Image File Format. Bildformat.

**UDK** - Universella Decimalklassifikationssystemet.

**UKOLN** - UK Office of Library Networking.

**URL** - Uniform Resource Locator. Syntax\* för webbadresser.

**URN** - Uniform Resource Number. Varaktigt identifikationsnummer för webbresurser.

**Webbsida** - Den information på en webbplats som kan nås utan att man klickar på en länk. Webbsidan motsvarar ofta det man kan se på skärmen eller det som tillgängliggörs när man använder rullningslistan.

**WWW (W3)** - World Wide Web eller "Webben". Det världsomspännande nätet. Ett nätverk av dokument som är kopplade till varandra via hyperlänkar. WWW är skapat av Tim Berners-Lee.

**W3C** - World Wide Web Consortium. Organisation som huvudsakligen ansvarar för utveckling av webbprotokoll.

**XML** - Extended Markup Language. Språk som ger möjligheter att skapa taggar som underlättar vid sökning.

**ZIG** - Z39.50 Implementors Group. Sammanslutningen som ansvarar för Z39.50-protokollets utveckling sedan 1990.

**Z39.50** - Ett protokoll\* för informationsöverföring i nätverksmiljö.

\* = står på annat ställe i ordlistan

## KÄLLFÖRTECKNING

### Elektroniska källor

Berners-Lee, Tim. *Metadata Architecture* . 1998.

<URL: <http://www.w3.org/Designissues/Metadata.html>> 981030

BIBLINK. <URL: <http://hosted.ukoln.ac.uk/biblink/>>981030

Bruun, Rasmus. *Er lykken gjort? - om metadata som indekseringsgrundlag ved bibliografisk kontrol af webdokumenter*. Danmarks Biblioteksskole. 1997.

<URL: <http://www.db.dk/student/k96/4/rb/ogaver/2semopg%5Ftekst.html>.> 981130

Caplan, Priscilla. "You Call It Corn, We Call It Syntax-Independent Metadata for Document-Like Objects." *The Public-Access Computer Systems Review* 6. 1995(4).

<URL: <http://info.lib.uh.edu/pr/v6/n4/capl6n4.html>> 981030

Danmeta. URL:<[http://nwi.dtv.dk/uinwi/index\\_e.html](http://nwi.dtv.dk/uinwi/index_e.html)> 981130

DC-1. <URL: <http://purl.org/dc/workshops/dc1conference/index.htm>>981230

DC-2. <URL: <http://www.oclc.org:5046/oclc/research/conferences/metadata2/>>981230

DC-3. <URL: <http://www.oclc.org:5046/oclc/research/conferences/imagemeta/>>981230

DC-4. <URL: <http://www.dstc.edu.au/DC4/>>981230

DC-5. <URL: <http://linnea.helsinki.fi/meta/DC5.html>>981230

DC-6. <URL: <http://purl.org/dc/workshops/dc6conference/index.htm>>990130

Dublin Core Metadata Initiative

<URL: [http://purl.oclc.org/metadata/dublin\\_core/main.html](http://purl.oclc.org/metadata/dublin_core/main.html)> 981130

Eriksson, Jörgen. *Metadata - the description of electronic objects, documents and service*. 1997.

<URL: <http://www.lub.lu.se/~jorgen/foredrag/meta.html>> 981030

Gradman, Stefan. "Cataloguing vs. Metadata: old wine in new bottles?" *64th IFLA General Conference August 16 - August 21, 1998*.

<URL: <http://www.nlc-bnc.ca/ifla/IV/ifla64/007-126e.htm>> 981030

Hakala, Juha. *Dublin Core in 1997 a report from Dublin Core metadata workshops 4 & 5*.

<URL: <http://linnea.helsinki.fi/meta/dcnord.html>> 981030

Hakala, Juha. Medförfattare: Hansen, Preben. Husby, Ole. Koch, Traugott och Thorborg, Susanne. *The Nordic metadata project Final report..* 1998.

<URL: <http://linnea.helsinki.fi/meta/nmfinal.htm>> 981130

Hakala, Juha. Husby, Ole och Koch, Traugott. *Warwick Framework and Dublin core set provide a comprehensive infrastructure for network resource description. Report from the Metadata Workshop II, Warwick, UK, April 1-3, 1996.*

<URL: <http://www.ub2.lu.se/tk/warwick.html>> 981030

Hedberg, Sten. *Nya vägar för boken.*

<URL: <http://www.kb.se/Nvb/>> 990430

Koch, Traugott. *Descriptions of the form and content of resources: Metadata.* 1996.

<URL: <http://www.ub2.lu.se/tk/metadata/metadata-general.html>> 981030

Lagoze, Carl. "The Warwick Framework - A Container Architecture for Diverse Sets of Metadata." *D-Lib Magazine*. Juli/Augusti 1996.

<URL: <http://www.dlib.org/dlib/july96/lagoze/07lagoze.html>> 981030

Library of Congress. *DISCUSSION PAPER NO: 99 Metadata, Dublin Core and USMARC : a review of current efforts, source:, date revised: 1997-01-21.*

<URL: [gopher://marvel.loc.gov:70/00/.listarch/usmarc/dp99.doc](http://gopher://marvel.loc.gov:70/00/.listarch/usmarc/dp99.doc)> 981030

Lynch, Clifford. "Searching the Internet, Combining the skills of the librarian and the computer scientist may help organize the anarchy of the Internet." *Scientific American*. 1997.

<URL: <http://www.sciam.com/0397issue/0397lynch.html>> 981030

Miller, Eric. "An Introduction to the Resource Description Framework." *D-Lib Magazine*. Maj 1998. <URL: <http://www.dlib.org/dlib/may98/miller/05miller.html>> 981030

Miller, Paul. "Dublin comes to Europe." *Ariadne*. 1998

<URL: <http://www.ariadne.ac.uk/issue14/dublin/intro.html>> 981030

Miller, Paul. *Metadata for the masses.* 1996.

<URL: <http://www.ariadne.ac.uk/issue5/metadata-masses/#search>> 981030

Nordic DC metadata creator. <URL: <http://www.lub.lu.se/cgi-bin/nmdc.pl>> 990130

Nordiskt Web Index. <URL: <http://nwi.ub2.lu.se/?lang=sv>> 990410

Rhyne, Charles. Chair of Art History at Reed College.

<URL: <http://www.dlib.org/dlib/january97/01weibel.html>> 981130

SAFARI. <URL: <http://www.hsv.se/verksamhet/utredningar/safari/>> 981230

SAFARI:s formulär. <URL: <http://safari.hsv.se/metadata/safari.pl?lang=sv>> 990130

SiL i Lund. <URL: <http://www.lu.se/lu/forskn/sil/>> 981230

Svenska datatermgruppen. <URL: <http://www.nada.kth.se/dataterm/>>

Svenska miljönätet. <URL: <http://smn.environ.se/>> 981230

Svenska miljönätets formulär. <URL: <http://smn.environ.se/miljonat/dc-gen/dcgen.htm>>990130

Svesök. <URL: <http://www.svesok.kb.se/>>981230

Svesöks formulär. <URL: <http://www.svesok.kb.se/user/submit/>> 990130

Swemeta.

<URL: <http://gungner.ub2.lu.se/cgi-bin/egwinit.pl/egwirtcl/metatargets.egw?swemeta>> 981130

Swick, Ralph. *W3C Technology and Society domain, Metadata Activity*. 1998.

<URL: <http://www.w3.org/Metadata/Activity.html>> 981130

Thiele, Harold. "The Dublin Core and Warwick Framework A Review of the Literature, March 1995 - September 1997." *D-Lib Magazine*. Januari 1998.

<URL: <http://mirrored.ukoln.ac.uk/lis-journals/dlib/dlib/dlib/january98/01thiele.html>> 981030

UKOLN. <URL: <http://www.ukoln.ac.uk/>> 981030

UKOLN Metadata Glossary. <URL: <http://www.ukoln.ac.uk/metadata/glossary.>> 981130

UKOLN (UK Office of Library Networking) Metadata Group. *An Overview of Resource Description Issues*. 1998.

<URL: [http://www.ukoln.ac.uk/metadata/desire/overview/rev\\_01.htm](http://www.ukoln.ac.uk/metadata/desire/overview/rev_01.htm)> 981030

Webguiden. <URL:<http://www.webguiden.dk.>> 981230

Weibel, Stuart. *Using Web Metadata The Dublin Core* . 1998.

<URL: <http://purl.oclc.org/~emiller/talks/www7/tutorial/part1/index.htm>> 981030

W3C The World Wide Web Consortium. <URL: <http://www.w3.org/>>

W3C Metadata Home Page. <URL: <http://www.w3.org/Metadata>> 981130

W3C RDF Home Page. <URL: <http://www.w3.org/Metadata/RDF>> 981230

## Tryckta källor

Cranfield, Andrew. "Z39.50 og danske biblioteker - en temadag: Danmarks Biblioteksskole, onsdag d. 3. december 1997". *DF-revy*. 21(2) 1998, s. 47f.

Hammar, Fredrik. *Nätverksbaserade informationssystem: Standardprotokoll ANSI Z39.50/OSI SR*. Högskolan i Borås. 1997.

Lange, Holley R. och Winkler, Jean B. "Taming the Internet: Metadata, a Work in Progress." *Advances in librarianship*. Vol 21, 1997. Academic press.

"Libraries for the New Millennium: Implications for managers." Red. David Raitt. *Library Association Publishing*. 1997. London.

Lingärde, Ulf. *Nya datorordboken*. 1998. Stockholm.

Tennant, Roy. "21st Century Cataloging". *Library Journal*. 123(7) 15 april 1998, s. 30.

Turner, Fay. "Selecting a Z39.50 Client or Web Gateway". *Library Hi Tech*. Nr 16(2) 1998, s. 7-18.

### **Otryckta källor**

E-postintervju med Annica Moshiri på Svenska miljönätet, december 1998.

E-postintervju med Lotta Jansson och Anette Svensson på Svesök, december 1998.

E-postintervju med Martin Sparr på SAFARI, december 1998.

E-postintervju med tre forskningsinstitutioner kopplade till SAFARI, februari 1999.

Hedberg, Sten. Jämförelse mellan SAFARI:s, Svenska miljönätets och Svesöks tillämpning av DC-metadata, december 1998. E-postkorrespondens januari - februari 1999.

Präntare, Åsa. Jämförelse mellan SAFARI:s och Svenska miljönätets tillämpning av DC-metadata, december 1998.