

DE PROFUNDIS

Varför behöver vi metadata?

Om vi nu har fått oss tilldelat inte mindre än *sju miljoner* för att bygga ett Svenskt miljönät, så antyder storleken av summan att statsmakterna anser att tjänsten bör ha en kvalité som går utöver vad befintliga tjänster kan erbjuda. Till exempel att man skall kunna söka på författare, ISBN-nummer, titel, en historisk tidsperiod eller geografisk plats som behandlas i ett arbete, osv. Jag uppfattar frågan om *vad* som skall vara sökbart som principiellt viktigt. Hur det skall göras är en teknisk fråga. Styrelsen bör vara medveten om att *om* det skall vara möjligt att göra sådana sökningar, *så* måste någon knappa in dessa data om data — *metadata*. Någon måste göra det jobbet under alla omständigheter. Man kan låta sekreterare göra det, eller låta bibliotekarier göra det. Det kan ske centralt hos en organisation hos Miljönätet, eller det kan ske hos medlemmarna. Det liggande förslaget innebär det sistnämnda.

Varför en standard?

Det är en fråga om en överenskommelse, eller som Tage Danielsson uttrycker det i förordet till sin *Grallimatik*:

Allvarligt talat: ord betyder ju ingenting egentligen. Varför heter hjärna hjärna till exempel? För att vi kommit överens om det. Vi skulle lika gärna kunna komma överens om att kalla det vi har i huvet förslagsvis *doja*. Exempel: *Albert Einstein har en av samtidens skarpaste dojos*. Eller "socialdemokratisk" skulle lika gärna kunna heta "lattjo" och "partiledare" "stare". Exempel: *Tage Erlander är en lattjo stare*. Skulle man kunna säga då.

Om vi nu har bestämt oss för att göra skäl för de pengar vi fått, så skulle vi kunna använda attributen "Auktor", "Aktstycke" och "Kvintessens" i våra databaser:

Attribut	Värde
Auktor	A.M. Lenngren
Aktstycke	Epigram
Kvintessens	<i>Att vitt och brett och lärt med kritisk recension en usel auktors verk i sanningsdagern ställa det är att fälla en tätting med kanon</i>

Det hade fungerat utmärkt. Det är bara det att ingen annan hade förstått vad våra poster egentligen betyder. Eller för att tala med Tage Danielsson:

Du och jag kan ju till exempel komma överens om att kalla en tur- och returbiljett Stockholm—Malmö för en åtsittande aftonklänning med svarta tofsar. Men jag lovar dig att vi får ett helvete i biljettluckan på Centralen.

Dublin Core är ägnad för precis det vi syftar till: förenklad beskrivning av elektroniskt material.

Inom Dublin Core har kunnigt folk, bibliotekarier, programmerare, arkivarier m fl diskuterat vilken uppsättning av attribut som med en minimal insats kan användas för att beskriva objekt på internet. Man har haft diskussioner med inriktning mot textdokument, bilder och multimedia. Och i december i fjol klubbade vi de femton attribut som nu föreligger som version 1.0 av standarden. Det är en inofficiell samförståndslösning, men officiella organ verkar rätta in sig i ledet. Library of Congress håller på att anpassa USMARC (MARC är det ledande katalogiseringsspråket i biblioteksvärlden) för att det skall gå att lätt omvandla DC poster till detta format. GILS försöker samordna med DC-initiativet. Vidare har det skrivits ett "internet draft" för DC 1.0, och utsikterna är goda att DC kommer att få status som Internetstandard. Till yttermera visso kommer DC sannolikt att anammas i många sammanhang även utanför Internetsfären. I dag sker förenklad katalogisering enligt icke standardiserade metoder av material som småtryck och liknande som väller in till arkivbiblioteken via pliktleveranser. DC erbjuder en standard för detta, som blir uppåt kompatibel med vad man gör med MARC-poster för andra media.

Varför bör medlemmarna stå för produktionen av *metadata*?

Jag uppfattar det som principiellt viktigt, och därmed en fråga för styrelsen att tillse att de *metadata* som produceras inom Miljönätet skall kunna användas på så många sätt som möjligt. Kodningsdetaljer är "bara" ett tekniskt problem, likaså hur dessa data bör samlas in. Det enda som är viktigt är att materialet kodas maskinläsbart, och på ett sätt som gör data återandvändningsbara.

Den föreslagna metoden innebär att medlemmarna har full kontroll över sin egen *metadataproduktion*. Den behöver alltså inte ske *enbart* för Miljönätet, utan potentiellt för andra sökmaskiner, och för intern användning inom respektive organisation.