

Innehållsdesign.
Verktyg, metoder och tillämpningar inom utbildningshistorisk
forskning och undervisning

Monica Langerth Zetterman

Uppsala september 2007

Utkast avhandlingsmanus inför seminarium den 9 oktober 2007.

Kommentator: Mats Dahlström, Bibliotekshögskolan, Borås.

Handledare: Donald Broady

Innehållsförteckning

Del I	Innehållsdesign	3
1	Inledning och syfte	5
1.1	Något om märkning av innehåll.....	6
1.2	Problembeskrivning	10
1.3	Pedagogisk relevans	16
1.4	Syfte och frågor	19
1.4.1	Avgränsning.....	20
1.5	Disposition	21
1.6	Något om terminologi	21
1.7	Om metod och material	25
2	Innehållsdesign för forskning och undervisning	31
2.1	Digitala arkiv för forskning och undervisning.....	34
2.2	Open Source – öppna standarder, metoder och verktyg.....	37
2.2.1	Öppen källkod och öppna program	37
2.2.2	Öppna standarder	40
3	Textorienterade metoder för innehållsdesign.....	45
3.1	Märkspråkens utveckling.....	48
3.2	COCOA	51
3.3	SGML – ett märkspråk för industrin	54
3.4	Innehållet och dess ordning.....	56
3.5	XML – en verktygslåda för webben	58
3.5.1	Migrering av textformat.....	61
3.5.2	Extern annotering.....	61
4	Dataorienterade metoder för innehållsdesign	65
4.1	Metadata för webben	70
5	Text Encoding Initiative.....	75
5.1	Riktlinjerna tar form.....	76
5.2	Från projekt till konsortium.....	79
5.3	Riktlinjernas utgångspunkter och principer	86
5.4	Modulerna.....	90
5.5	Från pizza till smörgåsbord	98

Del II	Tillämpad innehållsdesign	101
6	Ordningens praktik.....	103
6.1	Prosopografisk metod.....	103
6.2	Prosopografiska och utbildningshistoriska studier	105
6.3	Prosopografiska tillämpningar.....	105
6.4	Verktyg för prosopografiska studier.....	107
6.5	Prosopografisk innehållsmärkning	108
7	En kvinnohistorisk databas	113
7.1	Formering för offentlighet.....	113
7.2	Kvinnorna.....	116
7.3	Ordning och nyordning.....	118
7.4	Databasen i arbete.....	121
7.4.1	Webbarkivet.....	123
7.4.2	Interaktiva kartor	124
7.4.3	Nätverksanalyser	126
7.4.4	Beskrivande statistik	128
8	Kungl. Högre Lärarinneseminariet	131
8.1	Lärarinnestudenter vid Kungl. Högre Lärarinneseminariet.....	131
8.2	Digitaliseringsarbetet.....	133
8.3	Märkningsarbetet.....	133
8.4	Exempel på tillämpningar	134
9	Sammanfattande slutsatser och diskussion.....	135
10	Figur- och tabellförteckning.....	137
11	Bilagor.....	139
12	Referenser	189

Del I Innehållsdesign

1 Inledning och syfte

Utvecklingen av metoder och verktyg för internet medför många möjligheter för användning av digitalt material för forskning och undervisning. Dessa möjligheter nyttjas inte till fullo inom utbildningsvetenskaplig och utbildningshistorisk forskning och undervisning. Forskare, lärare och studerande är ofta beroende av hur digitalt material presenteras och framställs vid den webbplats som för tillfället besöks. Att göra något eget av innehållet brukar inte vara möjligt. Det saknas dock inte möjligheter. Tillgången av ändamålsenliga metoder och verktyg för att utforma digitalt innehåll så att materialet kan brukas och återanvändas på flera sätt.

Ett verksamhetsområde för utformning digital innehåll kan kallas för *innehållsdesign*.¹ I praktiken kan innehållsdesign innebära att beskriva, ordna och bearbeta digitalt innehåll med relevanta standarder och i plattformsberoende lagringsformat. Konkret handlar det om att med hjälp av märkning infoga information om själva innehållet – i detta sammanhang text – för att beskriva dess fysiska, logiska eller semantiska struktur. Med märkning av text skapas förutsättningar för att särskilja innehåll från form, något som senare nyttjas för vidare bearbetning och som medför att innehållet kan ordnas och omordnas, samt presenteras och användas på skilda sätt och i olika tillämpningar.²

Den här avhandlingen handlar alltså om innehållsdesign och om att utforska några av de metoder och verktyg som står till buds. För att utforska dess möjligheter utformar och använder jag tillämpningar av prosopografiskt material sammanställda för utbildnings- och socialhistoriska studier om kvinnor i offentlig verksamhet från 1880-talet fram till 1940-talet. Speciellt intresse riktas mot lärarinnor och de som studerade vid Kungliga Högre Lärarinneseminarier. Avhandlingen handlar dock inte om studier av kvinnors utbildning, yrkesverksamhet eller väg till offentlighet, utan om att skapa förutsättningar och möjligheter för en flexibel infrastruktur för utbildningshistorisk forskning och under-

¹ Begreppet innehållsdesign presenteras i Donald Broady, *Content design. Methods and tools for the creation of portable hypermedia archives. Notes for a proposed CID project. Version 2* (Nada/CID, KTH, 1997), s. 18-19, Donald Broady, "Digitala arkiv och portföljer," in *IT i skolan: mirakelmedicin eller sockerpiller? Observatoriet för IT lärande kunskap och kompetens, Rapport 45/2001* (Stockholm: IT-kommissionen, 2001).

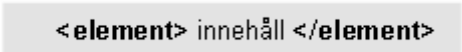
² Här avses till exempel olika slags analyser, transformering, extrahering, samt presentation till exempel publicering på webben.

visning, samt att utforska hur metoder och verktyg för innehållsdesign kan omsättas i praktiken.³

1.1 Något om märkning av innehåll

Innehållsdesign handlar som ovan nämnts om att beskriva, ordna, bearbeta och förbereda för presentation av innehåll. Nedan presenteras några inledande exempel på hur beskrivande märkning och ett exempel på tillämpningsområde. Inledningsvis nämndes att märkning av innehåll erbjuder möjligheter till att bruka ett och samma material på flera olika sätt. Men för att åstadkomma det krävs vissa förutsättningar, till exempel att materialet är noggrant och systematiskt beskrivet och strukturerat. Innan själva märkningen påbörjas behövs en modell som beskriver hur märkning av innehållets olika egenskaper skall gå till. Mer om det senare, först några exempel för den läsare som inte är bekant med märkning.

För att beskriva ett visst innehåll används märkord (element) som är inneslutna i avskiljare som markerar elementets början (<), benämning (**starttagg**) och slut (>), sammantaget blir det en **<starttagg>**. För att markera var det innehåll som skall inneslutas av märkningen sedan slutar används en specifik identifierare (*I*) som markerar att detta är en **</sluttagg>**.⁴ Ett element består av alltså av en "starttagg" och en "sluttagg" som markerar, beskriver och innesluter det valda innehållet.



```
<element> innehåll </element>
```

Figur 1. Exempel på den deskriptiva märkningens grundläggande principer: elementet.

Elementen kan sedan nästlas, det vill säga inneslutas i varandra, så att en hierarkisk struktur bildas som visas i figuren ovan.

³ Några resultat av statistiska analyser som gjorts på det prosopografiska materialet presenteras i Monica Langerth Zetterman, *Kvinnors mötesplatser och sociala och kulturella tillgångar kring sekelskiftet 1900. Beskrivning av prosopografiskt material inom projektet Formering för offentlighet*, 1 ed., vol. 41, *Rapporter från Forskningsgruppen för utbildnings- och kultursociologi* (Uppsala: Forskningsgruppen för utbildnings- och kultursociologi, Under utgivning).

⁴ För en ingående beskrivning av den deskriptiva märkningens grundläggande principer se, Charles F. Goldfarb and Yuri Rubinsky, *The SGML handbook* (Oxford: Clarendon, 1990), s. 5-17. samt kapitel 1 "A Gentle Introduction to XML" i Lou Burnard and C. M. Sperberg-McQueen, *Text Encoding Initiative. Guidelines for electronic text encoding and interchange, XML Version* (Oxford, Providence, Charlottesville, Bergen: Text Encoding Initiative Consortium, 2002).

```

<rot>
  <element1>Innehåll 1
    <element2>Innehåll 2
  </element2>
</element1>
</rot>

```

Figur 2. Exempel på den deskriptiva märkningens grundläggande principer: nästlade element.

Nedan ett annat exempel på hur märkningen kan se ut. Exemplet är hämtat ur det prosopografiska material som tillämpas i avhandlingen. Materialet är utformat och iordningställt för projektet *Formering för offentlighet. En kollektivbiografi över stockholmskvinnor 1880-1920*.⁵ I "Formering för offentlighet" har uppgifter samlats in om till exempel kvinnornas sociala bakgrund, utbildning, verksamhetsområden, yrkesliv, civilstatus, bostadsort, uppgifter om socialt umgänge, ekonomi, föreningsliv, utmärkelser och medlemskap i organisationer och andra sorters specifik symboliskt kapital.⁶

I figuren nedan beskrivs och ordnas ett urval av en persons egenskaper, i detta fall Ellen Key.

```

<node id="Ellen Key">
  <data id=N_23>
    <birthplace value="landsort">
      Född på landsorten
    </birthplace>
    <mentioned value="26">
      Omnämnt av 26 andra kvinnor
    </mentioned>
    <sphere value="PUB">Publicist </sphere>
    <education value="Lärarinneseminariet">
      Lärarinnekursen
    </education>
  </data>
</node>

```

Figur 3. Exempel på märkning av en kvinnas egenskaper och tillgångar.⁷

⁵ Donald Broady et al., "Formering för offentlighet. En kollektivbiografi över Stockholmskvinnor 1880-1920," in *Rapporter från Forskningsgruppen för utbildnings- och kultursociologi (Sociology of Education and Culture Research Reports)*, nr 22 (Uppsala: SEC/ILU, Uppsala universitet, 1998).

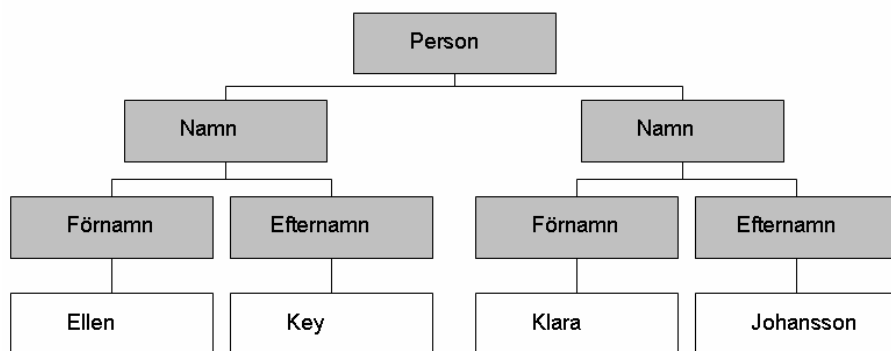
Se även <http://www.skeptron.ilu.uu.se/broadly/sec/ffo.htm> för en uppdaterad projektplan (2000), projektmedlemmar samt skrifter från projektet.

⁶ Se vidare avsnitt 6.4, s. 107 samt kapitel 7, s. 113.

⁷ Exemplet visar en förenklad märkning med GraphML där en hel del information är borttagen.

Denna beskrivning kan sedan användas och bearbetas på olika sätt. Man kan skapa index med hjälp av märkorden, sortera om ordningen av innehållet, utföra strukturerade sökningar, extrahera delar av det märkta innehållet eller infoga nya delar. Med hjälp av märkningen blir det även möjligt att flytta utvalda data till en databas. Vidare finns möjlighet att presentera innehållet i olika format, till exempel i HTML eller PDF-format. Det är även möjligt att migrera det märkta innehållet till andra lagringsformat och applikationer för att utföra olika slags analyser av materialet. Genom tillfoga en explicit beskrivning av innehållet ges även förutsättningar för andra att använda innehållet med den aktuella beskrivningen eller att kritisera, ändra och göra tillägg i vår beskrivning av innehållet.

Det vanligaste lagringsformatet för innehållsmärkning idag är XML, en förkortning som står för *Extensible Markup Language*. *Extensible* för att det kan utvidgas och anpassas samt *Markup Language* för att det använder märkord för avgränsning och strukturering av innehåll.⁸ I XML beskrivs och ordnas innehållet på det hierarkiska sätt som jag gett exempel på ovan, vilket innebär att innehållets struktur kan liknas vid ett träd. Detta kan med ett förenklat exempel se ut så här:



Figur 4. XML: hierarkisk dokumentmodell.

Deskriptiva metaspråk använder en hierarkisk dokumentmodell.⁹ I den modellen utgår man från att varje innehållsobjekt relaterar till, och ingår i, andra objekt. Ovanstående bild visar ett träd med sju noder. Det finns två typer av noder i detta träd och de kallas elementnoder respektive textnoder. De grå noderna är element och de vita noderna är den text, eller det innehåll, som finns i elementen **<förnamn>** och **<efternamn>**. Det måste alltid finnas ett, och endast

⁸ XML, *Extensible Markup Language*, är ett format för lagring och utbyte av data som är plattformsoberoende. XML är förenklat och härlett ur den mycket mer omfattande och äldre SGML-standard (ISO 8879:1986). Se vidare <http://www.w3.org/XML/>. Se även standardverket Goldfarb and Rubinsky, *The SGML handbook*.

⁹ I dessa sammanhang vanligen benämnd Ordered Hierarchy of Content Object (OHCO). Modellen presenteras och jämförs med andra dokumentmodeller bland annat i: Steven DeRose et al., "What is Text, Really?" *Journal of Computing in Higher Education* I, no. 2 (1990). Modellen har kritiserats av ett flertal forskare, se vidare avsnitt 5.3, s. 86.

ett, överordnat element i ett XML-dokument. Under roten, **<person>**, kan det dock i detta exempel finnas fler nästlade element, i detta fall är kvinnornas för och efternamn nästlade i ett överordnat namnelement. För att innehållet skall kunna beskrivas och struktureras samt reglera elementens nästling och relation till varandra enligt en viss dokumentmodell används ett schema som specificerar reglerna. Flera element kan förekomma på samma nivå enligt samma princip som de två "syskonelementen" **<namn>** i exemplet ovan. Hur elementen kan nästlas och/eller förekomma på samma hierarkiska nivå är beroende av dokumentmodellens (schemats) regler och restriktioner för den hierarkiska ordningen.

I strikt bemärkelse avser termen *element* en kombination av taggarna och det innehåll de innesluter och inte enbart starttaggen **<name>** och sluttaggen **</name>** som man kan föranledas att tro utifrån exemplet som presenterades i inledningen. I praktiken används dock *element* i båda betydelserna. Elementen är, som vi ovan sett, egentligen benämningen på de logiska delar innehållet byggs upp av: en bok består till exempel av kapitel som i sin tur innehåller stycken som består av meningar osv. Dessa element skiljer man åt genom att tillfoga markörer, till exempel en **<name>** tagg.¹⁰

Det är dock möjligt, om än inte önskvärt, att utforma en näst intill "platt" organisation av innehållet, med endast ett rotelement och med samtliga efterföljande element på nästa, och samma, nivå.¹¹ I den hierarkiska dokumentmodellen utgår man från att större innehållsobjekt innehåller mindre delar, som till exempel kapitel, stycken, meningar och ord, ned till minsta meningsbärande enhet.¹²

Den hierarkiska strukturen medför att innehållsobjekten och relationerna mellan dessa bestäms av en specifik ordning. Något som ibland kan vara en nackdel. Speciellt om man i ett och samma dokument behöver beskriva olika aspekter, som i sin tur har olika struktur jämfört med den ordnande principen. I den här studien ville vi bland annat beskriva kvinnornas kontaktnät. Då denna aspekt inte kunde inordnas i den dokumentmodell vi använde, löste jag det genom att ge varje annan kvinnas namn som nämndes i varje kvinnas biografi en unik identitet för att på så sätt relatera varje nämnt namn till en specifik kvinnas kontaktnät.¹³

I *Figur 5* nedan visas ett exempel på hur metoden kan användas. Här har märkningen använts som förutsättning för en nätverksanalys av de 101 kvinnornas kontakter inom gruppen.¹⁴

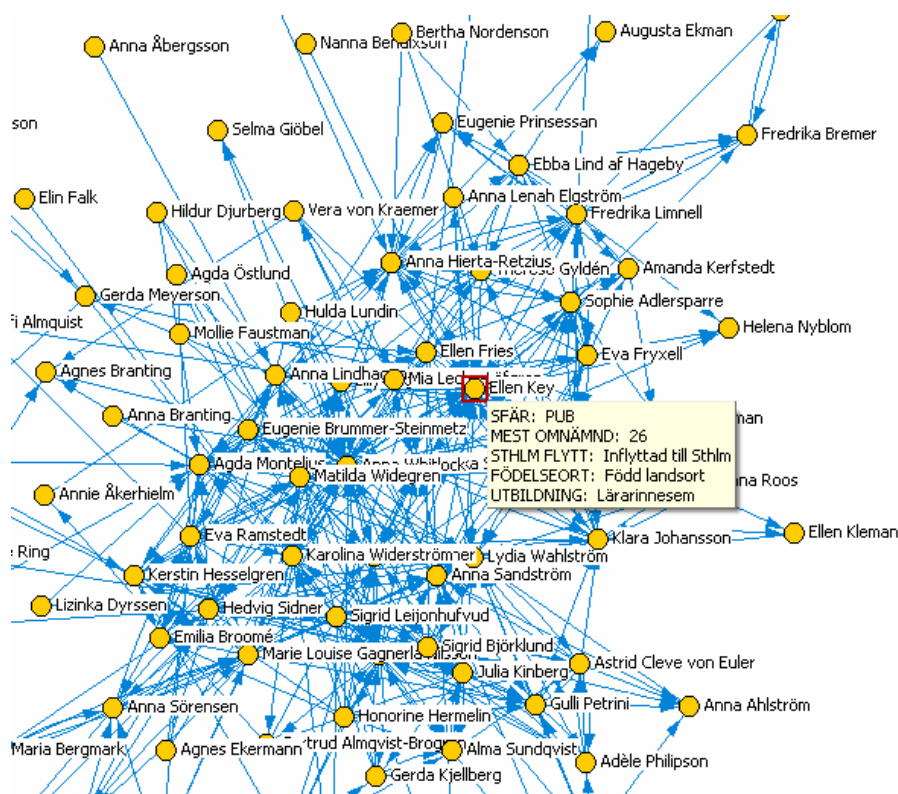
¹⁰ Goldfarb and Rubinsky, *The SGML handbook*, s. 18-19.

¹¹ Om dokumentmodellen utformas så, förloras också den flerhierarkiska ordningen av innehållet och texten representeras istället snarare i form av sekventiellt förekommande innehållsobjekt.

¹² Den minsta meningsbärande enheten varierar beroende på hur innehållet beskrivs. En lingvistisk analys beskriver oftast andra aspekter än en materialhistorisk analys.

¹³ Metoden för att beskriva kvinnornas kontaktnät presenteras närmare i avsnitt 7.4.3, s. 126 f.

¹⁴ Exemplet är ett urklipp, alla kvinnor är därför inte representerade i grafen. Se kapitel 7.4.3 Nätverksanalyser, s. 126 om nätverksanalyser.



Figur 5. Urklipp ur nätverksgraf med Ellen Keys egenskaper (attribut) accentuerade.

De kvinnor som hade många kontakter med övriga kvinnor i urvalet befinner sig i centrum av nätverksgrafen och kvinnorna med få kontakter återfinns således i periferin. Konstruktionen av graferna, det vill säga noderna och de riktade länkarna baseras på binära relationer, information som i sin tur utvunnits ur det märkta innehållet. Varje nod representerar en kvinna och i analysen kan man även välja att inkludera valfritt antal attributvärden som anger till exempel social bakgrund, utbildning, föreningsaktiviteter etc. Dessa värden kan sedan användas för filtrering och strukturering av graferna.¹⁵

1.2 Problembeskrivning

Användningar av digitala arkiv och andra datamängder har blivit en allt viktigare förutsättning, främst för forskning men även för undervisning. Inom samhällsvetenskaperna finns en tradition att bygga register av olika slag. Här an-

¹⁵ Se kapitel 7.4.3, s. 126 för exempel på filtrering och strukturering av grafer i nätverksanalyser.

vänds till exempel longitudinella databaser, individbaserade register eller register av specifika sociala fenomen.¹⁶ Humanister arbetar oftare med digitaliserat arkivmaterial, äldre tryck och språkkorporusar.¹⁷

Inom utbildningshistorisk forskning och undervisning kan man tänka sig att tillgång till mer komplexa och rikhaltigt innehållsmärkt material skulle kunna vara av intresse. Tillgången till digitala arkiv med historiskt källmaterial, klassiska utgåvor, vetenskapliga skrifter och textkorporusar är idag relativt god.¹⁸ Trots detta för de digitala arkiven en relativt undanskymd tillvaro som verktyg och hjälpmedel inom utbildningshistorisk forskning och utbildning i Sverige.¹⁹

I en kartläggning utförd av Vetenskapsrådet år 2005 gjordes en översikt av den nationella infrastrukturen inom samhällsvetenskap och humaniora, med fokus på material som insamlats eller sammanställts av forskare för forskning och som avses finnas tillgängligt för forskning nu och i framtiden. Slutsatserna av kartläggningen tyder på att materialen allt oftare utformas i samverkan mellan flera olika parter, att innehållet tenderar bli alltmer komplext samt att behovet av bättre tillgänglighet ökar. I rapporten föreslås en rad åtgärder, men först och främst betonas behovet av att:

[...] materialet dokumenteras och digitaliseras enligt gemensamma standarder (en förutsättning för integrering av olika material och komparationer).²⁰

I rapporten framhålls de humanistiska laboratorierna som särskilt viktiga för framtida resursutveckling.²¹ Därtill att verksamheten vid humanistlaboratorier med stor sannolikhet kommer att ha stor betydelse inom humanistisk forskning eftersom digitalisering, märkning och statistisk bearbetning spelar en allt större roll för forskning såväl som för undervisning. Det finns alltså all anledning till

¹⁶ Som exempel kan nämnas de register som används inom Forskningsgruppen för utbildnings- och kultursociologi (SEC) och som innehåller omfattande datamaterial från SCB över elever i svensk grund- och gymnasieskola, studerande i högskolan samt sociogeografiska data för den svenska befolkningen, för information om SEC:s forskningsprojekt där dataregistren används se: <http://www.skeptron.uu.se/broadly/sec/research.htm>, en översikt av dataregistren finns tillgänglig vid: <http://www.skeptron.uu.se/monicalz/esf/slides/Borjesson-Mikael.pdf>. Den demografiska databasen vid Umeå universitet är ett annat exempel på en omfattande forskningsdatabas, se <http://www.ddb.umu.se/>.

¹⁷ Vid Stockholms stadsarkiv finns ett antal databaser och register tillgängliga på webben, <http://www.ssa.stockholm.se/>. Vid den nationella arkivdatabasen finns det samlade nationella arkivbeståndet, <http://www.nad.ra.se/>. Stockholm Umeå Corpus (SUC) är en (morfosyntaktiskt) annoterad språkkorpus som utvecklats sedan 1992, se <http://spraakbanken.gu.se/parole/>.

¹⁸ För Sveriges del se exemplen i noten ovan, internationellt sett finns en mängd fritt tillgängliga av projekt och databaser inom samhällsvetenskap och humaniora.

¹⁹ [Belägg! Referens]

²⁰ Vetenskapsrådet, "Om forskningens infrastruktur inom humaniora och samhällsvetenskap i Sverige," (Vetenskapsrådet, 2005), s. 3.

²¹ Här framhålls HUMlab vid Umeå universitet där informationstekniken studeras, brukas och utvecklas. I rapporten nämns även Språk- och Litteraturcentrum vid Lunds universitet.

att studera vad vi kan lära av hur man inom humaniora utvecklat metoder och verktyg för att beskriva och ordna digitalt innehåll.²²

Vilken betydelse skulle innehållsdesign kunna ha för forskning och undervisning? Innehållsdesign skulle kunna medföra att ett och samma innehåll kan användas för flera olika syften och behov. En digital fulltextsamling eller databas som utvecklas för forskningsändamål skulle samtidigt kunna användas för undervisning genom att presentera endast valda delar eller sammanfoga det med nytt material utan att behöva ändra i det ursprungliga innehållet. Det skulle även finnas möjligheter till att utföra avancerade sökningar, kollationera, göra konkordanser, utvinna dokumentation om innehållet, samt tillföra nya annoteringar och kommentarer.²³

I rapporten från Vetenskapsrådet noteras även att humanistiska forskare ofta saknar kunskaper i bearbetning av forskningsmaterial.²⁴ Jag vet inte om förhållandet är detsamma inom samhällsvetenskaplig och utbildningsvetenskaplig forskning och undervisning. Med tanke på den ringa produktionen av tillgängliga digitalt material för forskning och undervisning i inom dessa områden (i Sverige) kan man anta att forskare, lärare och studenter inte känner till dessa möjligheter för hantering och utformning av plattformsberoende digitalt forsknings- och undervisningsmaterial. Istället används ofta proprietära lösningar som medför att det digitala innehållet inte kan flyttas, integreras eller återanvändas i andra sammanhang än vad som var tänkt från början.

Digitalisering är dessutom viktigt för att göra vårt gemensamma kulturarv tillgängligt, för allmänheten såväl som för forskning och utbildning. I en utredning från 2003 om Kungl. Bibliotekets verksamhet och arbetsformer, den så kallade KB-utredningen, inleds sammanfattningen med följande uppmaning:

Det är nu hög tid att göra en kraftsamling för informationsförsörjningen till forskning och utbildning.²⁵

I utredningen konstateras det även att flera länder i Europa²⁶ satsat stort på digitalisering av kulturarvet och att skälen till detta är många, till exempel av tillgänglighet och bevarandeskäl. För att Sverige skall kunna hävda sig internationellt är det viktigt att prioritera forsknings- och utbildningssektorn.

²² Se till exempel Text Encoding Initiative (TEI) som utvecklat de ledande riktlinjerna för beskrivning, lagring och utbyte av humanistiskt och samhällsvetenskapligt innehåll. Se även nedan, s. 15.

²³ I Litteraturbanken är det möjligt att tillfoga egna annoteringar i de digitala fulltextutgåvorna på webben, se <http://litteraturbanken.se/>

²⁴ Vetenskapsrådet, "Om forskningens infrastrukturer inom humaniora och samhällsvetenskap i Sverige.", ssk. s. 8.

²⁵ KB-utredningen, *KB - ett nav i kunskapssamhället: Kungl. biblioteket - Sveriges nationalbibliotek: verksamhet och visioner: betänkande, Statens offentliga utredningar, 2003:129* (Stockholm: Fritzes offentliga publikationer, 2003), s. 17.

²⁶ I utredningen (s. 177) nämns bl.a. Frankrike, <http://gallica.bnf.fr/> och Skottland, <http://scran.ac.uk/> [2007-03-23]

För att vi ska kunna ligga i täten måste staten satsa på att utveckla den digitala informationsförsörjningen till utbildning och forskning. Insatserna ska bidra till att regeringens mål om Sverige som ett informationssamhälle för alla uppfylls.²⁷

Utredarna tar på flera ställen upp vikten av *ordning och reda* och hur viktigt det är att använda gemensamma standarder för informationshantering. Även om KB-utredningen mer fokuserar på nationell samordning och bevarande av kulturarvet i ett vidare perspektiv än vad föreliggande studie gör, är resultaten i utredningen synnerligen relevanta för verksamhetsområdet innehållsdesign.

Digitalisering och standardisering är även föremål för utredning av EU-kommissionen. I en rekommendation från EU-kommissionen från augusti 2006 används begreppet *kulturellt innehåll* och även där betonas vikten av gemensamma standarder och normer för effektiv åtkomst och användning av digitalt innehåll.²⁸

En viktig aspekt vid digitalisering är alltså tillgänglighet, men även modularitet där märkningen, som en av andra möjliga metoder, kan medverka till att materialet blir utbyggbart, modifierbart och flyttbart så att andra kan bruka materialet utifrån egna önskemål:

Ett väl strukturerat och väl beskrivet innehåll behöver inte vara bundet till vissa presentationsformer eller undervisningsformer. Innehållet kan vara utspritt på nätet, det kan vara synnerligen heterogent och det behöver inte alls vara avsett enbart för undervisningsbruk. Om det är väl strukturerat och väl beskrivet kan lärare, studenter och elever överblicka och genomsöka det och sammanfoga material hämtat från olika håll.²⁹

Ett led i den riktningen är användningen av vedertagna märkningsscheman och riktlinjer för märkning av innehåll där innehållet, snarare än form är i centrum.³⁰ Tillgänglighet och flyttbarhet är i sig inget nytt eller revolutionerande, det formulerades bland annat för svensk del i forskningsprogrammet "Det nya handbiblioteket" för mer än tio år sedan.³¹ Det nya är snarare att nu, år 2007, är metoderna förbättrade och förenklade. Därtill är de verktyg som behövs relativt sett lättillgängliga jämfört med tidigare.

Vad vi kan se är alltså att trots att metoder och verktyg funnits tillgängliga i ganska många år har märkning av innehåll inte slagit igenom inom svensk ut-

²⁷ KB-utredningen, *KB - ett nav i kunskapssamhället: Kungl. biblioteket - Sveriges nationalbibliotek: verksamhet och visioner: betänkande*, s. 17.

²⁸ "Kulturellt innehåll - digitalisering, elektronisk tillgång och digitalt bevarande," ed. Europeiska gemenskapernas kommission (EU-kommissionen) (Europeiska unionens officiella tidning, 31 aug. 2006).

²⁹ Broady, "Digitala arkiv och portföljer," s. 11-16.

³⁰ Det bör betonas att märkning inte är den allenarådande metoden för att strukturera och ordna innehåll. De mest utvecklade och väl fungerande metoderna är relations- och objektdatabaser som passar bra för väl avgränsade och mer strikta modeller för innehållsmodellering än i detta fall. Mer om detta i kapitel 6, s. 103.

³¹ Broady, *Det nya handbiblioteket*, 1995 *Det nya handbiblioteket* finns tillgänglig vid: <http://www.skeptron.ilu.uu.se/broadby/dl/p-broadby-9304.htm>

bildningsvetenskaplig forskning eller undervisning i någon större omfattning. Varför kommer jag inte att försöka svara på i den här studien, men en försiktig spekulation är väl möjligen att metoder och verktyg ansetts som svårtillgängliga och att människor fortsätter göra det de är vana vid. Värdet av att investera tid och färdigheter på omständliga procedurer för att strukturera texter och annat innehåll kanske inte heller framstår som attraktiva lösningar när det finns färdiga lösningar att tillgå.

Dessutom skulle användare behöva olika typer av metoder och verktyg som är anpassade efter olika gruppers specifika behov och förkunskaper. Det som är obekant och krångligt kommer inte att efterfrågas. För att nyttja fördelar och möjligheter med innehållsdesign kanske det kan komma att behövs en bättre "beställarkompetens" hos lärare och studerande, därtill tillämpningar där metoder och verktyg provas. En lärare som vill använda digitala resurser i undervisningen behöver inte samma typ av verktyg som forskaren som själv producerar digitalt material har behov av. Vid vissa tillfällen behövs verktyg för att hantera och interagera med digitalt resurser, det vill säga sådana verktyg som innehållskonsumenter (till exempel lärare och studenter) kan behöva i undervisningen. Andra tillfällen kräver helt andra typer av verktyg. Sådant som forskare eller speciellt intresserade lärare och andra innehållsproducenter kan behöva för att beskriva, ordna och bearbeta digitalt material.

Saknas det vettiga användningsområden för innehållsdesign utbildning? Det gör det inte. Om innehållet i digitala arkiv var konstruerade utifrån en generell funktionalitet, som ett slags råämne, skulle forskare, lärare och studerande kunna använda ett och samma underlag för olika typer av användningar. Detta skulle även medföra möjligheter till en ny sorts digitala läromedel där funktion och användningsområden kan vara mer flexibla än i plattformsbärande digitala läromedel.³²

Dessutom skulle innehållsdesign kunna underlätta samarbete mellan lärare och i forskargrupper. Till exempel genom möjligheter att hålla reda på individuella bidrag i ett sammanfogat material genom att behålla uppgifter om ursprung och härkomst av de delar som tillförs till det större arkivet. Innehållsdesign innebär även möjligheter att sammanföra, beskriva och ordna ett material som sedan kan analyseras och bearbetas.³³ Eftersom ett sådant arkiv kan göras platt-

³² Till exempel XX vid XX.

³³ Att använda innehållsdesign och digitala arkiv för flexibelt forskningsmaterial har tidigare presenteras i Donald Broady and Monica Langerth Zetterman, "Formation for the public sphere – a collective biography" (paper presented at the DRH2003 (Digital Resources in the Humanities 2003), University of Gloucestershire, Cheltenham, England, 31 aug-3 sept 2003), Monica Langerth Zetterman, "Märkvärd. Noteringar om en prosopografi över kvinnor i offentlighet kring sekelskiftet 1900," in *Fältanteckningar: utbildnings- och kultursociologiska texter tillägnade Donald Broady*, ed. Mikael Börjesson, et al., *Skeptronserien*, (Uppsala: Forskningsgruppen för utbildnings- och kultursociologi ILU Uppsala universitet, 2006), Monica Langerth Zetterman, "Using Markup for Multivariate Analyses in the Prosopographical Study "Formation for the Public Sphere"." (paper presented at the ACH/ALLC Conference, Victoria University, Canada, June 15 - June 18 2005).

formsoberoende och modulariserat skulle det gå bra att flytta hela eller delar av arkivet, förutsatt att de upphovsrättsliga frågorna är lösta.³⁴

Forskning och utbildning är ju i många avseenden en textfokuserad praktik. Är digitala arkiv i utbildningssammanhang fortfarande förknippade med läsning snarare än användning? Det är möjligt. Digitala texter är ju dessutom inte särskilt läsvänliga att läsa på datorskärmen. Som ovan skisserats finns möjligheter att konstruera användarvänligt och flexibelt digitalt material. Till exempel med hjälp av avsedda metoder och verktyg om utvecklats inom textkritik, lingvistik och filologi, där skrifter och texter per se är studieobjekt och där det funnits ett reellt behov av ändamålsenliga verktyg.

Med uppgiften att beskriva och ordna ett visst innehåll följer nödvändigheten av att somligt inkluderas och accentueras samtidigt som annat dämpas eller exkluderas. Hur man klassificerar, beskriver och ordnar forskningsmaterial styrs av en rad olika faktorer, däribland forskningsområden. Beroende på hur ett innehåll beskrivs och ordnas ges också olika sorters möjligheter till användning. Givet det påverkas dessutom beskrivningen och ordningen dels av teoretiska utgångspunkter och antaganden med därav tillgängliga metoder och verktyg, dels av frågor av mer praktisk natur såsom källmaterialets kvalitet, forskningsgruppens sammansättning och kunskapsområden.

Problemet är också att det finns många verktyg och metoder att välja på. Vilka verktyg kan användas till vad? Vilka metoder passar till specifika användningsområden? För att försöka svara på dessa frågor kommer jag att granska två utvecklingslinjer som tillhandhåller och utvecklar metoder och verktyg för informationshantering och innehållsdesign. Dessa två utvecklingslinjer kan beskrivas som text- eller dokumentorienterade samt dataorienterade. De textorienterade metoderna är dominerande inom humaniora och samhällsvetenskap och de dataorienterade är dominerande inom data- och naturvetenskap. Inom de text- och dokumentorienterade utvecklingslinjen är det vanligaste verktyget Text Encoding Initiative³⁵ (TEI), den främsta de-facto standarden för märkning, lagring och utbyte av texter inom humaniora och samhällsvetenskap.³⁶ Den dataorienterade utvecklingslinjen tänker jag mig utgörs av den verktygslåda som den semantiska webben tillhandhåller.

³⁴ [Detta behöver jag diskutera någonstans]

³⁵ Text Encoding Initiative är en internationell organisation och en de-facto standard för digitala utgåvor av litterära och lingvistiska texter för forskning och undervisning inom humaniora och samhällsvetenskap samt för utbyte, lagring och bevarande av digitala texter. Se vidare <http://www.tei-c.org/>

³⁶ Den de-facto standard som nämns här består av dokument rekommendationer för hur märkning av specifika grupper av egenskaper bör märkas och i vilken ordning märkorden kan användas. Se kapitel 3 för diskussion om deskriptiv märkning, märkspråk och dokumentmodeller.

1.3 Pedagogisk relevans

På vilket sätt är den här avhandlingen pedagogiskt relevant? Nedan presenterar jag tre punkter som relaterar till den här studien och som jag anser har pedagogisk relevans. Den första är tillgängligheten, något som jag menar innefattar mer än tillgång till information. Förknippat med tillgänglighet, eller kanske svårtillgänglighet, är frågor om innehållets urval, funktionalitet och generalitet. Den andra punkten handlar om förutsättningar för att utforska det okända och det som inte är på förhand givet. Den tredje punkten är förutsättningar för användning av forskningsmaterial i undervisningen, det som i policydokument brukar benämnas forskningsanknytning.³⁷

I pedagogiska sammanhang är tillgång till information och kunskapsorganisation en nödvändig förutsättning då textbaserad information ofta utgör en stor del av undervisningen. Frågor man kan ställa sig är hur vi organiserar och framställer innehåll: vad som uppfattas som viktigt, hur vi kan ta till oss informationen eller vad vi kan använda innehållet till. Presenteras innehållet fragmentariskt eller i en kontext? Vad väljs ut? Vem gör urvalet?³⁸ Därtill har det bärande mediet betydelse för hur vi kan interagera med en text eller ett visst innehåll, en digitaliserad samling historiskt material ger andra förutsättningar för analyser och bearbetning än en tryckt bok. Texter brukar ses som antingen helt frikopplade från, eller intimt förknippade med, det bärande mediet. Espen Aarseth framhåller mediets betydelse i sin definition av text: en texts funktionalitet är att vidarebefordra verbal information. Av detta följer att innehållets funktionalitet är beroende av mediet och att text inte kan likställas med den information som förmedlas, förutsatt att information ses som teckensträngar som inte nödvändigtvis behöver vara meningsfull för betraktaren.³⁹

Mediet ramar alltså in, inte minst innehållets funktionalitet, genom de historiska, sociala och politiska sammanhang textens där de används men även tar över andra mediers funktioner och egenskaper genom att härma andra medietyper.⁴⁰ Webbsidors utformning liknar i många avseenden den tryckta bokens. Den hierarkiska dokumentmodellen som XML är ett exempel på är också den baserad på den tryckta textens design.

Texters funktion i forsknings- och undervisningssammanhang och strategier för en innehållslig infrastruktur som tillåter användare att arbeta med materialet

³⁷ Jfr Roger Säljö and Maria Södling, "Utbildning på vetenskaplig grund: röster från fältet," in *Högskoleverkets rapportserie, 2006:46 R* (Stockholm: Högskoleverket, 2006), s. 11.

³⁸ Veronica Johansson hävdar, i sin kunskapsöversikt och informationspolitisk analys av elektroniska dokument, att de aktörer och de maktintressen som påverkar informationshantering och tillgång relaterar till mer generella aspekter på tillgänglighet, tillgång och kunskapssyn i allmänhet. Veronica Johansson, "Elektroniska dokument i informationspolitisk belysning. En fråga om makt, kontroll och förhandlingar," *Human IT* 7, no. 2 (2004).

³⁹ Fritt efter Espen Aarseth, *Cybertext: perspectives on ergodic literature* (Bergen: Univ. of Bergen, 1995), s. 69-72. En intressant diskussion om digitala texters materialitet och medietypiska skillnader finns i: Mats Dahlström, *Under utgivning: den vetenskapliga utgivningens bibliografiska funktion, Skrif- ter från Valfrid, 34* (Borås: Valfrid, 2006), ssk. s. 70-76.

⁴⁰ Dahlström, *Under utgivning: den vetenskapliga utgivningens bibliografiska funktion*, s. 72 f. samt s.121.

hamnar inte alltid överst på dagordningen.⁴¹ En infrastruktur för forskning och undervisning med en bredare funktionalitet skulle kunna tillföra mer än att enbart presentera och tillgängliggöra olika slags digitalt material. Litteraturbankens digitala samlingar har en sådan uttalad målsättning. Den formulerar att, förutom presentation av digitala versioner av litterära verk, man även avser tillhandhålla verktyg som tillåter användare att arbeta med det digitala materialet.⁴²

Förutom möjligheterna till effektiva hjälpmedel för forskning och undervisning anser Svedjedal att tillgängligheten för studerande utanför universitets- och högskoleorterna såväl som för andra användargrupper är en viktig fråga:

Med tanke på de olika satsningar som nu görs på "nätuniversitet" bör Litteraturbanken också kunna bli en viktig faktor för de universitet som satsar på nya former av utbildning, ofta riktade till elever som studerar på mindre hemorter med begränsade biblioteksresurser. Samma nytta kan materialet naturligtvis göra för användare inom andra delar av utbildningssystemet: gymnasister och elever på folkhögskolor kan använda Litteraturbanken som underlag för uppsatser, redovisningar etc. Bakom detta ligger förhållandet att Litteraturbanken gör ett kulturellt viktigt textmaterial nationellt tillgängligt.⁴³

Tillgängligheten och nyttoaspekterna diskuteras även i KB-utredningen med skillnaden att där prioriteras tillgängliggörande av unika och svårtillgängliga verk för digitalisering, sådant som äldre tryck och handskrifter.⁴⁴

En annan aspekt på tillgänglighet är de digitala samlingar som publicerar verk som inte är svårtillgängliga på samma sätt som handskrifter och äldre tryck. Här tänker jag på opublicerade och oredigerade texter som snarare kan ses som arkivmaterial och som inte fallit inom ramen för de traditionella litterära verk som ingår i kurslitteraturen då de haft otillräcklig legitimitet (till exempel på grund av författarens kön).⁴⁵

I samhällsvetenskapliga och humanistiska ämnen är det inte alltid så enkelt att efterlikna ett laboratorium eller att göra experimentet. Digitala samlingar med tillgång till källmaterial och lämpliga verktyg kan ge studerande tillfälle att bearbeta och studera primärkällor och forskningsmaterial, söka information,

⁴¹ Här hänvisar jag till ovan anförda rapporter: KB-utredningen, *KB - ett nav i kunskapssambället: Kungl. biblioteket - Sveriges nationalbibliotek: verksamhet och visioner: betänkande*, Vetenskapsrådet, "Om forskningens infrastrukturer inom humaniora och samhällsvetenskap i Sverige."

⁴² I Johan Svedjedals sammanfattning i utredning om Litteraturbanken handlar en punkt tre om Litteraturbankens huvuduppgifter: *[...] samla in, arkivera och publicera material samt göra det möjligt för användare att arbeta med det*. Johan Svedjedal, "En svensk Litteraturbank? Utredning för Svenska Akademien, avlämnad i januari 2003," (2003), s. 42.

⁴³ *Ibid.*, s. 5.

⁴⁴ KB-utredningen, *KB - ett nav i kunskapssambället: Kungl. biblioteket - Sveriges nationalbibliotek: verksamhet och visioner: betänkande*, s. 213 f.

⁴⁵ En diskussion om maktförhållanden, tolkningsföreträde och idealiserade texter som måttstock för legitimitet av tidigare opublicerade litterära texter av kvinnor förs i: Julia Flanders, "Gender and the Electronic text," in *Electronic text: investigations in method and theory*, ed. Kathryn Sutherland (Oxford: Clarendon, 1997), ssk. s. 135 f.

tolka resultaten och formulera frågeställningar. Det vill säga genom att få tillfälle att praktisera forskningens hantverk.

Julia Flanders, vetenskaplig ledare för The Brown University Women Writers Project,⁴⁶ anser att en av de främsta pedagogiska utmaningarna med vetenskapliga digitala samlingar är möjligheterna att införa ett annat arbetssätt i undervisningen. Istället för ett komprimerat urval av texter, till exempel i en antologi, kan studenter ges möjlighet att konfronteras med excerpter såväl som hela verk, enskilda objekt såväl som kontexten och därtill bekanta sig med ett okänt innehåll. En utgångspunkt för detta, menar Flanders, är att studenterna ges möjlighet gå i närkamp med originaltexterna:

[...] huge, unmanageable texts; difficult texts; long, beautiful texts; absurd texts; poignantly misguided texts; fascinatingly dull texts; texts whose footnotes are the most interesting thing of all. The collection thus encourages students to think and act like their teachers, at least qualitatively – by asking the kinds of questions that only the full textual world can answer.⁴⁷

I Women Writers Project har man en uttalad målsättning att få lärandesituationer att likna forskarens arbetssätt och att, med viss hjälp, ta sig an ett *outförskat material*. Denna något idealistiska metod undgår dock inte ett visst mått av självkritik och pragmatisk realism. Ett digitalt arkiv med vetenskapliga utgåvor, versioner och textkritisk apparat kräver mycket arbete och resurser. Förutom det finns alltid risken att arkivet blir omodernt efter några år, då nya tekniker och metoder introduceras. Detta innebär att man istället valt en medelväg, mellan de stora universitetens generella digitala samlingar med många genrer och tidsperioder och en uppsjö av kringmaterial och mellan de kursspecifika webbaserade undervisningsmaterialen. Women Writers Project tillhandhåller därför bredvidmaterial till inskrivna texterna i form av kortare essäer om författarna och vissa ämnen.⁴⁸ Detta menar man hjälper studenterna att få en uppfattning om textens sociala sammanhang, historia och tillkomst.

En annan pedagogisk anknytning anser jag är att skapa förutsättningar för en forskningsanknuten högskoleundervisning genom användning av konkret forskningsmaterial i undervisningen. I en rapport om forskningsanknytning undervisning från Högskoleverket ges bland annat exempel på hur man vid olika lärosäten och utbildningar försöker införa en tydligare forskningsknuten

⁴⁶ The Brown University Women Writers Project (WWP), digitala samlingar och fulltextarkiv med texter av kvinnliga, ofta tidigare opublicerade, författare från perioden 1400-1850. Förutom uppbyggnad av digitala samlingar har man sedan tidigt 1990-tal utvecklat och utforskat verktyg, textteoretiska frågor, handledningar och "best practices" för textmärkning samt kursplaneutveckling. De digitala samlingarna är dock inte fritt tillgängliga utan kräver licens för åtkomst. Se <http://www.wwp.brown.edu/>.

⁴⁷ Julia Flanders, "Learning, Reading, and the Problem of Scale: Using Women Writers Online," *Pedagogy: Critical Approaches to Teaching Literature, Language, Composition, and Culture* 2, no. 1 (2002): s. 50.

⁴⁸ Se <http://www.wwp.brown.edu/about/index.html>.

undervisning med olika metoder: läsa och redovisa originalforskning, aktivt deltagande i seminarier, specifika och integrerade metodavsnitt med mera.⁴⁹ Rapporten avslutas med ett antal förslag till att stärka forskningsanknytningen i undervisningen.⁵⁰ Men inte i någon punkt diskuteras att man skulle kunna använda konkret forskningsmaterial såsom digitala samlingar eller databaser i undervisningen.

1.4 Syfte och frågor

Mot bakgrund av ovan skisserade möjligheter till infrastruktur för forskning och undervisning med innehållsdesign formuleras här det övergripande syftet för avhandlingen.

Mitt syfte är ett försök till metodutveckling genom att undersöka hur vi kan nyttja nya möjligheter till en infrastruktur för utbildningshistorisk forskning och undervisning med hjälp av innehållsdesign. Syftet försöker jag uppnå genom att jämföra utvecklingslinjer inom verksamhetsområdet innehållsdesign samt undersöka dess praktik genom att tillämpa metoder och verktyg för att beskriva, ordna och bearbeta digitalt innehåll.

Avhandlingen handlar inte om studier av kvinnors utbildning, yrkesverksamhet eller väg till offentlighet, utan om att skapa förutsättningar för utbildningshistorisk forskning och undervisning, därtill att utforska och ge exempel på hur metoder och verktyg för innehållsdesign kan omsättas i praktiken.

Följande frågor är vägledande för avhandlingens två delar:

Vilka särdrag kan urskiljas mellan den textorienterade och den dataorienterade utvecklingslinjerna?

På vilka sätt kan metoder och verktyg för att beskriva, ordna, bearbeta och presentera digitalt innehåll användas för utformning av utbildningshistoriska forskningsmaterial?

Hur kan innehållsdesignens metoder och verktyg tillämpas inom utbildningshistorisk forskning och undervisning?

Dessa frågor vägleder också avhandlingens disposition. Den första frågan används för att söka urskilja och jämföra utvecklingslinjer där metoder och verktyg för innehållsdesign utvecklas och tillhandahålls. Den andra frågan innebär att utforska de metoder och verktyg som kan användas för innehållsdesign.⁵¹

Den tredje frågan utforskas genom två exempel på tillämpningar. Tillämpningarna baseras på prosopografiskt material som utformats och iordningställts

⁴⁹ Säljö and Södling, "Utbildning på vetenskaplig grund: röster från fältet."

⁵⁰ Ibid., s. 59-62.

⁵¹ Detta avgränsas till verktyg och metoder som kan användas för att beskriva, ordna och manipulera digitalt innehåll för inom humanvetenskaperna. Metoder för innehållsdesign inom naturvetenskap och matematik berörs inte av studien.

inom projektet *Formering för offentlighet. En kollektivbiografi över stockholmskvinnor 1880-1920* där borgerliga kvinnors vägar till offentligheten studeras och dels i en studie om lärarinnestudenter. Den första tillämpningen är en prosopografisk databas över 1250 offentligt verksamma kvinnor hämtade ur samtida biografiska uppslagsverk⁵² och detaljerade kapitalbeskrivningar (biografiska beskrivningar) om 101 kvinnor.⁵³ Den andra tillämpningen är en prosopografisk databas med uppgifter om samtliga (1980 kvinnor) inskrivna lärarinnestudenter vid Kungl. Högre Lärarinneseminariet under hela verksamhetsperioden 1861-194. De två prosopografiska databaserna sträcker sig över en relativt lång historisk period, en tid när borgerliga kvinnor började få möjligheter till högre utbildning och att träda ut i offentliga sammanhang.

1.4.1 Avgränsning

Avhandlingen består av två delar, en första del där jag beskriver och analyserar metoder och verktyg, något som i sin tur utgör en förutsättning för del två. Där används ett urval av verktygen till två liknande, men skilda prosopografiska databaser. Ramen för avhandlingen är det prosopografiska materialet tillsammans med de metoder och verktyg som används för att beskriva, ordna och presentera textbaserat innehåll. Detta gör att själva innehållet, eller källmaterialet, utgör en avgränsande faktor. De verktyg jag använder en annan och slutligen den prosopografiska metoden en tredje. Därav följer att avhandlingen är mer praktikorienterad och resultatfokuserad än andra pedagogiska studier. Resultaten, det vill säga själva databaserna och därtill relaterat material kommer att publiceras på webben och vara fritt tillgängliga att användas.

Avsikten med min studie är alltså att utforska en uppsättning verktyg och metoder och att iordningställa utbildningshistoriskt material för användning i forskning och undervisning inom framför allt utbildningshistoria. Därför har jag inte haft för avsikt att studera undervisning på webben. Ej heller lärande i webbaserade utbildningssituationer eller kunskapens sociala dimension i den forskningspraktik jag deltagit i under utformning och uppbyggnaden av det prosopografiska materialet. Jag är förstas medveten om att kunskapsprocessen inte är individuell eller historielös, utan ett resultat av ett kollektivt arbete. Det kollektiva arbete och de beslut som medverkat till databasernas utformning presenteras och diskuteras i anslutning till aktuellt avsnitt. Ett experiment inte kan förstås utan hjälp av ett tankekollektiv menar Fleck. Det enstaka experi-

⁵² Uppgifterna är avskrifter ur framför allt Nils S. Lundström, *Svenska kvinnor i offentlig verksamhet: porträtt och biografier*, Uppsala 1924 men även Valborg Hedberg & Louise Arosenius, *Svenska kvinnor från skilda verksamhetsområden: biografisk uppslagsbok*, Bonniers, Stockholm 1914.

⁵³ Kapitalbeskrivningarna är författade av: Donald Broady, Boel Englund, Joans Gustavsson, Lena Hammarlund, Ingrid Heyman, Lena Kårelund, Agneta Linné, Eva Trotzig, Kerstin Skog-Östlin. Se även förteckning i [[bilaga XX]]

mentet ges därför en begränsad betydelse relativt den sammansatta erfarenheten.⁵⁴

Studien har en interdisciplinär utgångspunkt, något som känns nödvändig av tämligen pragmatiska skäl. Den praktik som är en förutsättning för avhandlingens första del och för dess tillämpningar är framför allt ett resultat av forskning och utvecklingen inom områden som biblioteks- och informationsvetenskap, datavetenskap, filologi, litteraturvetenskap, språkvetenskap och historia.

Jag gör även anspråk på att ge ett bidrag till utbildningshistoriska studier. Här tillkommer alltså ytterligare ett interdisciplinärt område. I inventering av utbildningshistorisk forskning poängteras att:

Utbildningshistoria är inte att betrakta som en vetenskaplig disciplin i den mening att det konstituerar ett eget forskarutbildningsämne eller för den delen rymms inom ett enda ämne. Detta faktum har gjort att denna inventering inte utgått från disciplinräns. I stället har forskningens innehåll fått styra. Som gemensam nämnare för den forskning som inkluderats i denna inventering är därför det historiska studiet av bildning, utbildning, undervisning, fostrande och lärande. Historiska studier implicerar här inte heller någon absolut tidsgräns utan inkluderar även nutids- och samtidshistoria.⁵⁵

Risken att jag griper efter för mycket och därmed ger mig ut på tunn is är inte obetydlig. Jag försöker dock avgränsa mig till att presentera och använda mig av sådant som direkt, eller åtminstone indirekt, relaterar till utformning och märkning av digitalt material i allmänhet och prosopografiskt material i synnerhet.

1.5 Disposition

[detta skrivs senare]

1.6 Något om terminologi

I avhandlingen förekommer en del termer som kan behöva definieras. Vissa av termerna har en mer specifik teknisk innebörd än andra och för dessa hänvisas läsaren till anförda noter och referenser. Några termer är dock centrala för innehållet och de diskuteras nedan. Inledningsvis har jag talat om metoder och verktyg, innehåll och form, märkning av innehåll samt digitala material. Vissa begreppen kanske kan tyckas självklara och just därför kommer jag att tydliggöra innebörden av dem i relation till den här studien.

⁵⁴ Ludwik Fleck, *Uppkomsten och utvecklingen av ett vetenskapligt faktum: inledning till läran om tankestil och tankekollektiv*, trans. Bengt Liliequist (Eslöv: B. Östlings bokförl. Symposion, 1997), s. 102 f.

⁵⁵ Esbjörn Larsson, *Att studera titlar som kapital: en teoretisk diskussion kring militära titlar i 1700-talets svenska samhälle* (2003), s. 7.

Innehållsdesign är som sagt centralt för den här studien och med innehållsdesign avses ett verksamhetsområde där man använder sig av speciellt ämnade metoder och verktyg för att utforma digitalt innehåll. Med hjälp av innehållsdesign söker man särskilja mellan det intellektuella och semantiska innehållet och vad som är form, struktur och ordning. Med innehåll avses här det stoff, tankemängd⁵⁶ eller ämne som framförs och presenteras i skriftlig eller muntlig form – i motsats till innehållets yttre utformning och presentation. Termen innehåll avser även de uppfattningar och betydelser som kan inrymmas i ett begrepp. Detta är en särskilt viktig aspekt då märkningen erbjuder möjligheter att explicit tillskriva ett visst begrepp en eller flera specifika betydelser. I avhandlingen används innehåll synonymt med text särskilt i de sammanhang då jag inte talar om avhandlingens material. Det material som används som studieobjekt är visserligen textbaserat men merparten består av registerposter och inte av prosaliknande texter. En stor del av materialet består dessutom av digitala bilder. Därför används i första hand termen innehåll, förutom där det är särskilt befogat eller passande att använda text. Nämnas bör att de verktyg och metoder jag använder är konstruerade med utgångspunkt i att beskriva och strukturera textbaserad information. En ganska oproblematisks utgångspunkt kan man tycka. Peter Shillingsburg påpekar dock att en text förstås har olika innebörder beroende på vem du frågar. För vissa forskare är texten det fysiska objektet, andra anser att text är serier av tecken eller symboler medan vissa menar att text är abstrakta begrepp.⁵⁷

I vissa avsnitt används termen resurs relativt flitigt och termen material används genomgående i avhandlingen. Jag tänker mig en skillnad som består i att en resurs är något, en viss form av innehåll, som är tillverkat, färdigställt eller åtminstone sammanställt och som är tillgängligt för användning och bearbetning. Man skulle också kunna säga att en digital resurs är en sorts produkt och därmed resultatet av en designprocess. Till skillnad från ett material som är utgångspunkten i denna designprocess, det vill säga det innehåll som används som utgångspunkt för att konstruera något nytt eller något annat. Ett material behöver heller inte vara en representation av en originalutgåva eller ett helt nyproducerat innehåll, utan kan bestå av en sammanställning eller omorganisering av andra resurser. Givetvis är gränserna ofta flytande mellan vad som utgör en resurs och ett material, kriteriet för material är dock att det används som utgångspunkt för att skapa något annat.

Jag har ofta undrat varför man ibland använder elektronisk och ibland digital när man menar samma sak? Det talas om digitala bibliotek, elektronisk publice-

⁵⁶ Från fornsvenskan, jfr tyska Inhalt. Källa Nationalencyklopedin
http://www.ne.se/jsp/search/article.jsp?i_art_id=O194337

⁵⁷ Peter L. Shillingsburg, *From Gutenberg to google: electronic representations of literary texts* (Cambridge: Cambridge Univ. Press, 2006), s. 12.

ring, elektroniska texter, e-text, e-post etc. Å ena sida kan man hävda att det inte så viktigt med ett stringent termbruk, huvudsaken är att människor vet vad som avses. Å andra sidan skulle en hopblandning av termer kunna bidra till osäkerhet om vad som är vad. Vad skiljer digitala texter från elektroniska dito?

Först något om skillnaden mellan begreppen elektronisk och digital. Med digital anges värden med hjälp av sifferuttryck – oftast i form av det binära talsystemet⁵⁸ – i motsats till analog där värden anges som punkter på en skala.⁵⁹ Betydelsen av elektronisk avser sådant som är eller nyttjar elektronik. Den största skillnaden är alltså inte mellan elektronisk och digital, utan mellan digital och analog. I praktiken handlar det om vilken lagringsmetod som används. Datorer behöver både elektronik och digital teknik för att fungera, en DVD behöver elektronik för att användas, däremot inte för att behålla den lagrade informationen. Information lagrad med hjälp av analog teknik behöver också ofta elektronik, kassettband lagrar ljud och data analogt och magnetiskt men behöver elektroniska verktyg för att användas.

Susan Hockey, en av föregångarna inom *Humanities Computing*, publicerade för några år sedan en översikt och historisk genomgång av elektroniska texter inom humaniora. Hockey framhåller att skillnaden mellan elektroniska texter och digitala bilder framför allt ligger i funktionaliteten. Hon menar att digitala texter är avbildningar av materiella föremål, till exempel en digital faksimilutgåva, som inte tillåter manipulering utan syftar till att vara en exakt representation av en originalutgåva.⁶⁰ Poängen med att särskilja elektroniska texter från digitala texter skulle då vara att de elektroniska texterna väsentligen förbättrar möjligheterna att manipulera innehållet jämfört med digitala avbilder som lagras i till exempel PDF-format.⁶¹ Mot bakgrund i diskussionen ovan om skillnader mellan digital och analog anser jag att man inte bör skilja elektronisk och digital på det sätt som Hockey gör då dessa tekniker är varandras förutsättningar.⁶² Digital teknik är grunden för datorers funktionalitet och används oavsett om jag producerar PDF-filer eller om jag transkriberar en text i ett ordbehandlingsprogram. Elektroniska texter grundas på digital teknik och skillnaden i graden av manipulerbarhet beror på i vilket digitalt format innehållet lagras. Lagras innehållet med hjälp av PDF-formatet finns mindre möjligheter till senare manipulation, dock ger formatet en möjlighet till versionshantering och långtidsbeva-

⁵⁸ Det binära talsystemet är ett positionssystem med basen två i motsats till vårt vanliga decimalsystem som har tio som bas. I det binära systemet används vanligen siffrorna 0 eller 1 för att beteckna positioner och de decimala talen betecknas baserat på potenser av 2. Till exempel talet 45 betecknas 101101 som då betecknar: $32+0+8+4+1$. Siffrorna 0 och 1 kallas ofta för bitar, från engelskans *bit* (binary digit).

⁵⁹ I matematisk betydelse är analoga värden kontinuerliga till skillnad mot digitala värden som är diskreta.

⁶⁰ Detta ger förstås en större säkerhet i fråga om närhet till originalets visuella egenskaper.

⁶¹ Se Susan M. Hockey, *Electronic texts in the humanities: principles and practice* (Oxford: Oxford University Press, 2000). s. 1 ff.

⁶² Detta är giltigt utom för själva lagringen, där behövs ingen elektronik. På en hårddisk lagras data på en skiva med magnetiskt material och bevaras utan elektronik, för att skriva in och läsa det data som lagrats på hårddisken behövs dock elektronik.

rande av digitala dokument.⁶³ Om innehållet däremot lagras i XML-formatet blir det möjligt att senare välja format för presentation, som då kan vara PDF, HTML etc.

Bruket av elektronisk text, vill jag hävda, härstammar från den tidiga datoranvändningen då det blev praxis att säga e-post, e-lärande, e-texter etcetera. Termen digital används oftare i sammanhang som specifikt rör hur data representeras och lagras. Information, eller innehåll, lagras digitalt och kan sedan presenteras som bilder, texter eller grafik.

Jag försöker därför genomgående använda prefixet *digital* när jag talar om texter, bilder eller annat material som produceras, lagras och manipuleras med hjälp av datorer och tillhörande applikationer. Vid de tillfällen när jag berör innehållsdesign i relation till det för innehållsdesign dominerande anglosaxiska verksamhetsfältet *Humanities Computing* används ibland en direkt översättning av den vanligt förekommande termen *electronic text*.

I titeln används termerna metoder och verktyg: Vad är skillnaden mellan metoder och verktyg för innehållsdesign? Skillnaden är inte helt klar. Vissa metoder har utvecklats och förenklats genom olika applikationer vilket medför att de senare tillskrivits verktygens egenskaper. Ett verktyg beskrivs ofta i konkreta ordalag som ett redskap som kan användas "för att bearbeta material"⁶⁴ och anses ofta fungera som vår förlängda arm. En metod har däremot en mer övergripande betydelse och kan beskrivas som ett "planmässigt tillvägagångssätt för att uppnå ett resultat"⁶⁵. Hur ett verktyg utformas och brukas kan sägas vara beroende av både vilken metod som används och både av materialets beskaffenhet. Ett verktyg är därmed mer än bara ett verktyg, då det står i relation till användning, sammanhang, intentioner och bakgrund.⁶⁶

En metod i det här sammanhanget kan då vara principer och idéer om hur text kan beskrivas och ordnas, genom att infoga märkord i texten för att beskriva en godtycklig mängd egenskaper av ett visst innehåll. SGML är ett metaspråk och därmed en metod som föreslår hur systematisk beskrivning och ordnande av textinnehåll kan gå till. Ett exempel på tvetydigheten är XML som är en delmängd av metoden SGML och utformades som ett verktyg för att beskriva och utbyta innehåll på webben till skillnad från den mycket mer komplexa SGML-specifikationen. Varken SGML eller XML säger något om till exempel vilka element som bör användas för att beskriva ett godtyckligt innehåll, istället

⁶³ En metod för långtidsbevarande och unika ID:nr av digitala publikationer har implementerats bland annat i DiVA systemet, utvecklat vid Uppsala universitet. Se t.ex. Eva Müller et al., "The DiVA Project - Development of an Electronic Publishing System," *D-Lib Magazine* 9, no. 11 (2003).

⁶⁴ Nationalencyklopedin: uppslagsord "verktyg".

⁶⁵ Nationalencyklopedin: uppslagsord "metod".

⁶⁶ Willard McCarty argumenterar för vikten av att granska de verktyg vi använder, framför allt de epistemologiska grunderna och skillnaden mellan de "verkliga" objekten (*real-world objects*) och det sätt vi beskriver och manipulerar objekten. Willard McCarty, *Humanities computing* (Basingstoke; New York: Palgrave Macmillan, 2005), s. 37-41.

anges generella regler för hur ett välformulerat dokument formatet måste se ut. Detta medför, anser jag, att varken SGML eller XML kan anses vara verktyg. Det faktum att SGML och XML specifikationen anger generella regler för dokumentets struktur är viktigt då en rad olika verktyg (applikationer) bygger på dessa principer.

Som exempel kan nämnas att på det sätt som webbläsare tolkar och visar innehåll är beroende av den hierarkiska trädstruktur som SGML:s riktlinjer bygger på.⁶⁷ Webbläsarna är dock programmerade att vara överseende med felaktigheter i dokumentstrukturen (HTML-koden) och innehållet kommer att visas även om dokumentet har fel struktur.⁶⁸ HTML är däremot inte en metod, utan en applikation – en dokumenttypsdefinition – av SGML. Enligt denna logik är även TEI:s riktlinjer också ett verktyg – av flera andra möjliga – som tillämpar metoderna SGML och XML för att beskriva och ordna innehåll.

Till sist vill jag nämna att forskning och praktik kring märkning har skett huvudsakligen i den anglosaxiska forskningsvärlden. Min uppfattning är att den svenska terminologin saknar en tydlig och stringent begreppsanvändning. Ett sådant exempel är ordet märkning, som i vissa sammanhang benämns kodning, annotering eller uppmärkning. I engelskan används *encoding* och *markup* ofta synonymt. *Encoding* (kodning) är den process som gör att information kan transformeras till ett annat format. Till exempel på det sätt tecken kodas digitalt genom en binär notation. En text är alltså alltid kodad enligt något format utan att för den skull innehålla explicit märkning.⁶⁹ *Markup* används i praktiken ofta med liknande betydelse som *encoding*, trots att det är ett substantiv, men avser snarare själva aktiviteten att tillföra koder till den redan tecken-kodade texten. Märkning (*markup*) innebär alltid att man kodar något, men kodning (*encoding*) behöver inte innehålla märkning.⁷⁰

1.7 Om metod och material

Tiden runt sekelskiftet 1900 erbjuder många möjligheter för studier av kvinnors väg till offentlighet, kamp för utbildning, deras verksamheter, inträde på arbetsmarknaden eller kamp för rösträtt. Den här studien handlar dock inte primärt om dåtidens borgerliga kvinnors väg till offentlighet. Däremot om att skapa förutsättningar för sådana studier. För att undersöka hur vi kan gå tillväga

⁶⁷ Mer om trädstruktur i kapitel 3, s. 45 f.

⁶⁸ Den nya standarden för webbdokument, XHTML, är dock inte lika förlåtande för felaktiga innehållsstrukturer och kommer att "tvinga" innehållsproducenter att skriva välformulerade webbdokument.

⁶⁹ Textens tecken måste representeras av ett specifikt schema, för vår del Latin-1 (ISO-8859) eller det numera vanligare Unicode formatet som klarar av att representera många fler tecken än tidigare kodscheman. Därtill är en text alltid kodad i och med styckesindelningar, punkter, mellanrum, versaler etc.

⁷⁰ Se även HUMANIST listans arkiv (Vol. 14) för diskussion om kodning (*encoding*) vs märkning (*markup*), http://lists.village.virginia.edu/lists_archive/Humanist/v14/subject.html.

för att digitalisera och sammanställa uppgifter hämtade ur olika slags källor samt ordna, bearbeta och använda dess uppgifter använder jag mig av ett explorativt arbetssätt för att utforma prosopografiskt material.

Arbetsättet är explorativt i den bemärkelsen att jag kombinerar metoder för digitalisering med socialhistoriska och samhällsvetenskapliga metoder för att studera hur denna kombination kan användas för utbildningshistoriska studier.

I syftet finns en underliggande ambition att inte bara beskriva och utforska ett antal metoder utan även att, om möjligt, generera nya frågor om hur vi kan utforma och använda digitalt material för forskning och undervisning. I detta ingår också ett bakomliggande instrumentellt syfte med tanke på att tillämpningarna, genom utforskandet, förhoppningsvis kan sammanföra kunskap från flera områden till en informerad forsknings- och undervisningspraktik. Målsättningen är identifiera metoder eller problem som närmare kan behöva utforskas och därmed skapa en grund för fortsatta studier. Studien är även beskrivande då jag strävar efter att kartlägga och redogöra för metodernas karaktäristika genom att inte bara systematisera och beskriva strukturen, utan även tillämpa, dels metoder och verktyg, dels de modeller som används för utformning av det prosopografiska materialet. Studien är mer orienterad mot nya frågor än mot svar och mer mot *att* modellera snarare än *om* modellering. Dokumentmodeller för beskrivning av, och modellen för insamlingen av det prosopografiska materialet, utforskas alltså genom att sätta verktygen och metoderna i arbete. Willard McCarty beskriver modelleringens tre roller och relationer mellan objekt och subjekt:

[the] manipulatory essence of modelling, with its connotation of embodied action, physically or metaphorically; the mediating role and ternary relationship modelling establish between knower and known; the directed vector-like engagement of the inquirer's attention, *through* the model he or she has made *to* the object of study; and the model's consequent function as an artificial agent of perception and instrument of thought.⁷¹

Vad framträder genom modellen och vad får stå tillbaka i ett empiriskt material? En modells fokuserade och förenklande funktionalitet ger oss möjligheter att fokusera det vi letar efter men även risken att finna det vi inte söker efter.⁷²

Det empiriska materialet består av olika slags texter. Dels texter om de metoder och verktyg som används för utformningen av det digitala material som handlar om kvinnorna. Dels de biografiska texter och de arkivposter som innehåller uppgifter om de kvinnor som ingår i detta prosopografiska digitala material.

⁷¹ McCarty, *Humanities computing*, s. 38.

⁷² Som när modellen inte passar materialet – eller tvärtom. En diskussion om hur det överskott som modellen förkastat utmanar oss i användandet av modellen, metoderna och verktygen i avsnitt **Error! Reference source not found.**, s. **Error! Bookmark not defined.** f. Se även Ibid., s. 37-41.

Det empiriska materialet består i del I av artiklar, riktlinjer, diskussionslistor, medlemslistor och metainformation om texter. I del II består det empiriska materialet av en blandning digitaliserade och inskrivna primär- och sekundärkällor samt biografier.⁷³ Det empiriska materialet är alltså tämligen disparat.

Den prosopografiska metoden som närmare diskuteras i kapitel 6 består i att systematiskt samla in ett stort antal uppgifter om en avgränsad grupp människor som har något gemensamt till exempel kön eller en viss socialgrupp. I det här sammanhanget är det dels borgerliga kvinnor kring sekelskiftet 1900 och dels lärarinnestuderande under perioden 1861-1943 som avgränsar in gruppen. John Tosh belyser nödvändigheten av en eklektisk strategi för historiska studier av sociala grupper:

Sociala grupper efterlämnar inga kollektiva arkiv. Deras sammansättning och plats i den sociala strukturen måste rekonstrueras ur ett brett spektrum av källor som har satts samman av högst banalare skäl.⁷⁴

Den här studien baseras just på ett sådant arbetssätt, källmaterialet härrör från olika typer av källor och metoderna för insamling och sammansättning varierar. Avhandlingen handlar dessutom om klassificeringsproblematiken och andra ställningstaganden som är nödvändiga att göra när vi konstruerar historiskt material om sociala grupper. Metoden att sammanställa samma sorts uppgifter för alla individer som kan sägas tillhöra en avgränsad grupp kallas för *prosopografi*, även kallad kollektivbiografi.

Till uppgiften att strukturera och sammanställa ett visst innehåll, i det här fallet prosopografiska databaser, tillkommer även aktiviteten att klassificera, något som kommer problematiseras och uppmärksammas i avhandlingen.

Modellering av innehållet inbegriper även klassificering som tjänar två huvudsakliga syften där båda två är lika viktiga: att separera och att gruppera. Resultatet av klassificeringen synliggör därmed, i bästa fall, de ställningstaganden och de utgångspunkter som utgör de kategoriserade objektens karaktäristika.⁷⁵ En grundläggande uppgift för klassificeringar är alltså att på ett systematiskt sätt indela och organisera fenomen eller objekt⁷⁶, vilket samtidigt innebär vissa prioriteringar och urvalsförfaranden. Detta kan ske utifrån explicita eller implicita teorier kombinerat med pragmatiska överväganden. Ett exempel på en välkänd klassificeringspraktik är biblioteksvärldens klassifikationssystem. Under slutet av 1800-talet skapades bibliotekens universella klassifikationssystem som grundades på föreställningar om att klassificeringen och dess struktur i en objektiv

⁷³ Materialet redovisas mer detaljerat i kapitel 6, s. 103 och **Error! Reference source not found.**, s. **Error! Bookmark not defined.** samt i Bilaga 7.

⁷⁴ John Tosh, *Historisk teori och metod*, 2., rev. uppl. ed. (Lund: Studentlitteratur, 2000), s. 114.

⁷⁵ C. M. Sperberg-McQueen, "Classification and its Structures," in *A companion to digital humanities*, ed. Susan Schreibman, Raymond George Siemens, and John Unsworth, *Blackwell companions to literature and culture*; 26 (Blackwell Pub., 2004).

⁷⁶ Objekt är den tekniska term som används i klassificeringsteori. Ibid.

mening korresponderade mot kunskapsområdets egen struktur.⁷⁷ Detta synsätt ifrågasätts från många håll inte minst inom vetenskapssociologin.⁷⁸ I den här undersökningen rör det sig om att klassificera och konstruera objekt med utgångspunkt i mer konkreta sociala företeelser som utbildning, socialt ursprung och individers verksamhetsområden.

Att beskriva och att ordna genom att klassificera är i praktiken ofta föremål för återkommande förhandlingar, omförhandlingar och förändringar. En av de grundläggande principerna för klassificering är just nödvändigheten av att somligt innehåll inkluderas och accentueras samtidigt som annat dämpas eller exkluderas. Begreppet klassificering syftar alltså såväl till *resultat* såväl som själva verksamheten *att klassificera*. Hur man klassificerar, beskriver och ordnar forskningsmaterial styrs av en rad olika faktorer, däribland forskningsområden. Beror på hur ett visst innehåll beskrivs och ordnas ges också olika möjligheter till användning, det vill säga bearbetning och presentation. Givet det påverkas dessutom beskrivningen och ordningen dels av teoretiska utgångspunkter med därav tillgängliga metoder och verktyg, dels av frågor av mer praktisk natur såsom källmaterialets kvalitet, forskningsgruppens sammansättning och kunskapsområden.

Pierre Bourdieu formulerar i *Homo Academicus* objektens konstruktionsprinciper och säger att vi ska låta objekten tala för sig själva genom att systematiskt blottlägga de implicita strukturer och klassificeringar som "redan finns i *själva objektet*".⁷⁹ Denna så kallade objektivisering innebär dels en brytning med våra vardagsföreställningar om objektet och dels en brytning med vardagsspråket. För att bryta vår vardagliga förståelse behöver vi rekonstruera alla indikatorer som förutsätter det studerade fenomenet. Dessutom behöver vi vaska fram de relevanta egenskaper⁸⁰ inom det specifika fält där dessa egenskaper har någon betydelse - det vill säga de kapitalformer som har något värde inom det fält som studeras.⁸¹ Det relationella synsättet på konstruktioner av fakta är framträdande hos Ludwik Fleck⁸² som menar att fakta inte existerar i någon objektiv mening utan är en produkt av våra vardagsföreställningar som påverkas av tankestilens

⁷⁷ Se Karin Pettersson, *Den manliga heterosexuella normen: en analys av SAB-systemet utifrån ett feministiskt och queerteoretiskt perspektiv*, Magisteruppsats i biblioteks- och informationsvetenskap vid Bibliotekshögskolan/Biblioteks- och informationsvetenskap, 2001:73 (Borås: Högsk. i Borås Bibliotekshögskolan/Biblioteks- och informationsvetenskap, 2001).

⁷⁸ Se t.ex. Pierre Bourdieu, *Homo academicus*, trans. Lars Eberhard Nyman and Mats Rosengren, *Moderna franska tänkare*, 27 (Eslöv: Brutus Östlings bokförl. Symposion, 1996), Karin D. Knorr-Cetina, *Epistemic cultures: how the sciences make knowledge* (Cambridge, Mass.: Harvard University Press, 1999), Andrew Pickering, *Science as practice and culture* (Chicago: Univ. of Chicago Press, 1992).

⁷⁹ Pierre Bourdieu, *Homo Academicus* 1996, p. 43, (kursivering i original).

⁸⁰ Vad som är av betydelse och därmed relevant utgör förutsättningarna för en vetenskaplig undersökning. Auguste Comte, Émile Durkheim, and Max Weber, *Tre klassiska texter* (Göteborg: Korpen, 1991), s. 124.

⁸¹ Bourdieu, *Homo academicus*, s. 44.

⁸² Fleck, *Uppkomsten och utvecklingen av ett vetenskapligt faktum: inledning till läran om tankestil och tankekollektiv*.

tvång som i sin tur är grundad i vår medverkan i olika tankekollektiv. En kunskapsprocess är att skapa användbara resultat – faktum – ur givna förutsättningar, det vill säga den kunskap vi redan har. Resultaten påverkas av ett tvång – tankestilen – som har en stiliserande effekt och som uppfattas av oss som den objektiva verkligheten. De fenomen, eller de kategorier vi skapar genom klassificeringar, är produkter av rum såväl som tid. Något som har betydelse för hur vi tolkar de egenskaper och fenomen vi sedan försöker beskriva. Att ha bostad på Östermalm vid tiden runt förra sekelskiftet innebar inte alls samma status som idag. Att som kvinna studera på universitetet för 100 år sedan var en kamp och förbehållet några få, att kvinnor dessutom inte kunde få statlig anställning efter utbildningen påverkade förstås också vissa familjers vilja att bekosta utbildning för sina döttrar.⁸³

I den här studien handlar klassificeringen konkret om att beskriva, ordna och särskilja uppgifter om gemensamma karaktäristika och tillgångar för individerna i det prosopografiska materialet. Särskilt intresse riktas mot skillnader i konstruktioner av vetenskapliga analytiska kategorier av sociala företeelser genom användning av antingen (externt) etablerade klassificeringar eller internt konstruerade klassificeringar avsedda att spegla detta specifika innehålls interna logik och ordning.

Den modell som används i detta sammanhang är den hierarkiska dokumentmodellens en första utgångspunkt då de metaspråk som här presenteras baseras på just denna modell.⁸⁴ En hierarkisk dokumentmodell reflekterar den lingvistiska uppfattningen om att text består av sekventiellt ordnade objekt. I den hierarkiska dokumentmodellen är begreppet innehållsobjekten av primär betydelse och de är av olika slag såsom (ord, meningar, stycken, citat etc.). Samma uppsättning innehållsobjekt framställas och formas på olika sätt något som också påverkar betydelsen.⁸⁵ De flesta innehållsobjekten ingår i större objekt, en mening är en del av ett stycke som är en del av ett avsnitt etc. Denna, i huvudsak abstrakta textmodell representerar ett synsätt på text som medel för överföring av information, består av lingvistiska sekvenser och är hierarkiskt uppbyggda.⁸⁶

Espen Aarseths modell skiljer sig från den hierarkiska dokumentmodellen genom att han ger de materiella förutsättningarna och de funktionella effekterna större betydelse i sin pragmatiska och breda definition av text:

⁸³ Se t.ex. Tord Rönholm, *Kunskapens kvinnor: sekelskiftets studentskor i mötet med den manliga universitetsvärlden*, *Forskningsrapporter från Institutionen för historiska studier vid Umeå universitet*, 12 (Umeå: Institutionen för historiska studier Univ. distributör, 1999). för en studie om det kvinnliga studenten i en manlig miljö, deras sociala kapital, bakgrund och framtid kring förra sekelskiftet.

⁸⁴ Ordered Hierarchy of Content Objects, OHCO. Se även s. 8 f.

⁸⁵ DeRose et al., "What is Text, Really?"

⁸⁶ Allen H. Renear, "Out of Praxis: Three (Meta)Theories of Textuality," in *Electronic text: investigations in method and theory*, ed. Kathryn Sutherland (Oxford: Clarendon, 1997). Kritik mot den hierarkiska dokumentmodellen, se kapitel 5.3, s. 86 f.

At text, the, is any object with the primary function to relay verbal information. Two observations follow from this definition: 1) a text cannot operate independently of some material medium, and this influences its behavior; 2) a text is not equal to the information it transmits.⁸⁷

I definitionen särskiljer Aarseth mellan text och information, där informationen består av teckensträngar med två modaliteter. Dels teckensträngarna såsom de framträder för läsarna (*scriptons*), dels såsom de är arrangerade i texten (*textons*)⁸⁸. Därtill har text en transversal funktion, det vill säga den mekanism som framställer och ordnar informationen, det vill säga *scriptons*, genererade ur *textons*, till en text.⁸⁹ I en statisk text är *scriptons* konstanta, i en dynamisk text kan de ändras medan *textons* antingen är konstanta eller deras antal och innehåll kan ändras. Jag anser att Aarseths modell är användbar för att förstå idén bakom skillnaden mellan en texts innehåll och form men även för att den ger mediet, och därmed innehållets materiella förutsättningar, men även graden av manipulation (mänsklig eller maskinbaserad) en avgörande roll.⁹⁰ Betydelsen av mediet för en texts framställning och funktionalitet har inte alltid varit självklar. Inom till exempel lingvistik har en immateriell tolkning av text varit dominerande sedan länge. Den modellen har visat sig vara svår att överföra på digitala texter och som en reaktion på det har andra modeller för textbegreppet arbetats fram av olika forskare.⁹¹

⁸⁷ Aarseth, *Cybertext: perspectives on ergodic literature*, s. 70.

⁸⁸ *Textons* såsom de grundläggande elementen för uppbyggnaden såsom själva byggnadsställningen.

⁸⁹ Aarseth, *Cybertext: perspectives on ergodic literature*, s. 70.

⁹⁰ En statisk webbsida med länkar till olika innehållsobjekt har ett fixt antal *scriptons* och *textons*, medan ett datorspel eller en databas kan generera olika uppsättningar *scriptons* med utgångspunkt i en uppsättning *textons*. I en chat som tillåter inmatning av text är antalet ingående *textons* förstås okänt. Och vid en dynamisk webbplats, som t.ex. en wiki som tillåter inmatning av text, är *textons* såväl som *scriptons* föränderliga.

⁹¹ Här hänvisar till Mats Dahlströms redogörelse om textbegreppets materiella betingelser. Dahlström, *Under utgivning: den vetenskapliga utgivningens bibliografiska funktion*. Se även Johan Svedjedals framställning om olika typer av textbärare och dess olika förutsättningar för strövtåg i texten. Johan Svedjedal, *Den sista boken* (Stockholm: Wahlström & Widstrand, 2001), s. 11-32.

12 Referenser

- Aarseth, Espen. *Cybertext: perspectives on ergodic literature*. Bergen: Univ. of Bergen, 1995.
- André, Jacques, Vincent Quint, Richard Furuta, and Institut national de recherche en informatique et en automatique. *Structured documents, The Cambridge series on electronic publishing*. Cambridge: Cambridge University Press, 1989.
- Barnard, D, L. Burnard, J.P. Gaspart, L.A. Price, C. M. Sperberg-McQueen, and G.V. Varile, eds. *Hierarchical Encoding of Text. Technical Problems and SGML Solutions*. Edited by Nancy Ide and Jean Vâeronis. Vol. 29, *Computers and the Humanities, The Text Encoding Initiative: Background and Contents*, 1995.
- Berners-Lee, Tim, and Mark Fischetti. *Weaving the Web: the original design and ultimate destiny of the World Wide Web by its inventor*. 1. ed. San Francisco: Harper, 1999.
- Berners-Lee, Tim, James Hendler, and Ora Lassila. "The Semantic Web." *Scientific American*, no. 279.5 (2001).
- Bourdieu, Pierre. *Homo academicus*. Translated by Lars Eberhard Nyman and Mats Rosengren, *Moderna franska tänkare*, 27. Eslöv: Brutus Östlings bokförl. Symposion, 1996.
- Bradley, John. "Finding a Middle Ground between 'Determinism' and 'Aesthetic Indeterminacy': a Model for Text Analysis Tools." *Literary and Linguistic Computing* 18, no. 2 (2003).
- Bradley, John, and Harold Short. "Using Formal Structures to Create Complex Relationships." In *Resourcing Sources*, 3-21. Oxford: Unit for Prosopographical Research, 2002.
- Bradley, John, and Paul Vetch. "Supporting Annotation as a Scholarly Tool - Experiences From the Online Chopin variorum Edition." *Literary and Linguistic Computing* 22, no. 2 (2007): 225-41.
- Broadly, Donald. *Content design. Methods and tools for the creation of portable hypermedia archives. Notes for a proposed CID project. Version 2: Nada/CID, KTH*, 1997.
- . "Det nya handbiblioteket." In *Biblioteken, Kulturen och den sociala intelligensen. Aktuell forskning inom biblioteks- och informationsvetenskap*, edited by Lars Höglund, 83-107. Göteborg: Forskningsrådsnämnden/Valfrid, 1995.
- . "Digitala arkiv och portföljer." In *IT i skolan: mirakelmedicin eller sockerpiller?* 67. Stockholm: IT-kommissionen, 2001.
- . "French prosopography: definition and suggested readings." *Poetics* 30 (2002): 381-85.

- . "Hur kodar man Röda Rummet?" In *ABM, IT och forskningen. Rapport från en konferens på Kungliga Biblioteket den 17 november 1999*, edited by Mats Rolén. Stockholm: Riksbankens Jubileumsfond, 2000.
- Broady, Donald, Boel Englund, Ingrid Heyman, Agneta Linné, Kerstin Skog-Östlin, Eva Trotzig, and Annika Ullman. "Formering för offentlighet. En kollektivbiografi över Stockholmskvinnor 1880-1920." In *Rapporter från Forskningsgruppen för utbildnings- och kultursociologi (Sociology of Education and Culture Research Reports)*, nr 22. Uppsala: SEC/ILU, Uppsala universitet, 1998.
- Broady, Donald, and Hasse Haitto. "Internet and the humanities: the promises of Integrated Open Hypermedia." Paper presented at the Contemporary computer and network technologies, Moskva, 17-18 Jan. 1996.
- Broady, Donald, and Monica Langerth Zetterman. "Formation for the public sphere – a collective biography." Paper presented at the DRH2003 (Digital Resources in the Humanities 2003), University of Gloucestershire, Cheltenham, England, 31 aug-3 sept 2003.
- Brockman, W. S., L. Neumann, C. L. Palmer, and T. J. Tidline. "Scholarly Work in the Humanities and the Evolving Information Environment." Washington, D.C.: Digital Library Federation, Council on Library and Information Resources, 2001.
- Burnard, L. "New tricks from an old dog: An overview of TEI P5." Paper presented at the Dagstuhl Seminar Proceedings. Digital Historical Corpora - ARchitecture, Annotation, and Retrieval, Dagstuhl, Germany 2007.
- . "Report of Workshop on Text Encoding Guidelines." *Literary and Linguistic Computing* 3 (1988): 131-33.
- Burnard, L., and Sebastian Rahtz. "RelaxNG with Son of ODD." Paper presented at the Extreme Markup Languages 2004, Montréal, Québec 2004.
- Burnard, Lou, Katherine O'Brien O'Keeffe, and John Unsworth. *Electronic textual editing*. New York, N.Y.: Modern Language Association of America, 2006.
- Burnard, Lou, and C. M. Sperberg-McQueen. *Text Encoding Initiative. Guidelines for electronic text encoding and interchange, XML Version*. Oxford, Providence, Charlottesville, Bergen: Text Encoding Initiative Consortium, 2002.
- Börjesson, Mikael. "Om att definiera sociala eliter och att undersöka deras utbildningsstrategier." Paper presented at the Svenska historikermötet 2005, Uppsala universitet, 22-24 april 2005 2005.
- Cameron, Averil, ed. *Fifty years of prosopography: the later Roman Empire, Byzantium and beyond*. Edited by Averil Cameron, *Proceedings of the British Academy*, 118. Oxford: Published for the British Academy by Oxford University Press, 2003.
- Carlson, G. "Literary Works in Machine-Readable Form." *Computers and the Humanities* 1 (1967): 75-102.
- Cederblad, Carl. *Bildningens väg: idéerna och livet*. Uppsala: Lindblad, 1932.

- Chartier, Roger. *Böckernas ordning: läsare, författare och bibliotek i Europa mellan 1300-tal och 1700-tal*. Translated by Jan Stolpe. 1. uppl. ed. Göteborg: Anamma, 1995.
- . *Forms and meanings: texts, performances, and audiences from codex to computer, New cultural studies*. Philadelphia: Univ. of Pennsylvania Press, 1995.
- Comte, Auguste, Émile Durkheim, and Max Weber. *Tre klassiska texter*. Göteborg: Korpen, 1991.
- Coombs, J.H., Allen H. Renear, and Steven DeRose. "Markup systems and the future of scholarly text processing." *Communications of the ACM* 30, no. 11 (1987): 933 - 47.
- Cummings, James, and Monica Langerth Zetterman. "Legacy Data Migration: A pilot study on the methodological feasibility of conversion and enhancement of electronic resources." In *Joint International Conference of the Association for Literary and Linguistic Computing and the Association for Computers and the Humanities*. Gothenburg, Sweden, 2004.
- Dahlström, Mats. "Nya medier, gamla verktyg." *Human IT* 6, no. 4 (2002): 77-116.
- . *Under utgivning: den vetenskapliga utgivningens bibliografiska funktion, Skrifter från Valfrid, 34*. Borås: Valfrid, 2006.
- DeRose, Steven. "Markup Overlap: A review and a Horse." Paper presented at the Extreme Markup Languages, Montréal, Québec, August 2-6, 2004 2004.
- DeRose, Steven, David G. Durand, Ellen Mylonas, and Allen H. Renear. "What is Text, Really?" *Journal of Computing in Higher Education* I, no. 2 (1990): 3-26.
- DeRose, Steven J. *The SGML FAQ book: understanding the foundation of HTML and XML, Electronic publishing series*. Boston: Kluwer Academic, 1997.
- Driscoll, M.J. "Levels of Transcription." In *Electronic textual editing*, edited by L. Burnard, Katherine O'Brien O'Keefe and John Unsworth, 254-61. New York, N.Y.: Modern Language Association of America, 2006.
- . "XML markup of prosopographical data." Paper presented at the Exploring New Methods for Prosopography in the Humanities and the Social Sciences, Uppsala, Sweden 2007.
- Eck, Werner. "The *Prosopographia Imperii Romani* and Prosopographical Method." In *Fifty years of prosopography: the later Roman Empire, Byzantium and beyond*, edited by Averil Cameron, 11-22. Oxford: Published for the British Academy by Oxford University Press, 2003.
- Engman, Jonas. "Handling fine-meshed filtering from archive to resource level. Combining RDF-metadata with XML content markup." Umeå University, 2004.
- "European Interoperability Framework for pan-European eGovernment Services." In *EIF Publication*. Luxembourg: Office for Official Publications of the European Communities, 2004.
- Fevrell, Valter. *Kungl. Högre lärarinneseminariet in memoriam: minnesrunor*. Stockholm: Geber, 1943.

- Flanders, Julia. "Gender and the Electronic text." In *Electronic text: investigations in method and theory*, edited by Kathryn Sutherland, 127-43. Oxford: Clarendon, 1997.
- . "Learning, Reading, and the Problem of Scale: Using Women Writers Online." *Pedagogy: Critical Approaches to Teaching Literature, Language, Composition, and Culture* 2, no. 1 (2002): 49-59.
- Fleck, Ludwik. *Uppkomsten och utvecklingen av ett vetenskapligt faktum: inledning till läran om tankestil och tankekollektiv*. Translated by Bengt Liliequist. Eslöv: B. Östlings bokförl. Symposion, 1997.
- Florin, Christina. *Kampen om katedern: feminiserings- och professionaliseringsprocessen inom den svenska folkskolans lärarkår 1860-1906, Umeå studies in the humanities*, 82. Umeå & Stockholm: Almqvist & Wiksell, 1987.
- Fähraeus, Rudolf. *Högre lärarinneseminariets historia*. Stockholm: Norstedt, 1943.
- Garshol, Lars Marius. "Living with topic maps and RDF." Paper presented at the XML Europe 2003, London, May 5-8 2003.
- Giordano, Richard. "The TEI header and the Documentation of Electronic Texts." In *Text encoding initiative: background and context*, edited by Nancy Ide and Jean Vâeronis, 75-98. Dordrecht; London: Kluwer Academic, 1995.
- Goldfarb, Charles F. "SGML: The Reason Why and the First Published Hint." *Journal of the American Society for Information Science* 48, no. 7 (1997).
- Goldfarb, Charles F., and Yuri Rubinsky. *The SGML handbook*. Oxford: Clarendon, 1990.
- Heckscher, Ebba. *Några drag ur den svenska flickskolans historia: under fleres medverkan samlade*. Stockholm: Norstedt & Söner, 1914.
- Hedberg, Valborg, and Louise Arosenius. *Svenska kvinnor från skilda verksamhetsområden: biografisk uppslagsbok*. Stockholm: Bonnier, 1914.
- Heyman, Ingrid. *Gänge hatt till: omvårdnadsforskningens framväxt i Sverige: sjuksköterskors avhandlingar 1974-1991, Skeptronserien*. Göteborg: Daidalos, 1995.
- Hillyar, Anna, and Jane McDermid. *Revolutionary women in Russia, 1870-1917: a study in collective biography*. Manchester: Manchester University Press, 2000.
- Hockey, Susan M. *Electronic texts in the humanities: principles and practice*. Oxford: Oxford University Press, 2000.
- Huitfeldt, C. "Multi-Dimensional Texts in a One-Dimensional Medium." *Computers and the Humanities* 28 (1995): 235-41.
- Ide, Nancy, and C. M. Sperberg-McQueen. "The TEI: History, Goals, and Future." In *Text encoding initiative: background and context*, 5-15. Dordrecht; London: Kluwer Academic, 1995.
- Ide, Nancy, and Jean Vâeronis. *Text encoding initiative: background and context*. Dordrecht; London: Kluwer Academic, 1995.
- Initiative, Text Encoding. "The TEI Consortium: A Membership Prospectus." 2001.
- Johansson, Veronica. "Elektroniska dokument i informationspolitisk belysning. En fråga om makt, kontroll och förhandlingar." *Human IT* 7, no. 2 (2004): 197-264.

- Johns, Adrian. *The nature of the book: print and knowledge in the making*. Chicago: Univ. of Chicago Press, 1998.
- Kay, M. "Report on an Informal Meeting on Standard Formats for Machine-readable Text." In *Literary Data Processing Conference 1964 Proceedings*, edited by Jess B. Bessinger, 327-28. N.Y.: White Plains, 1965.
- KB-utredningen. *KB - ett nav i kunskapssamhället: Kungl. biblioteket - Sveriges nationalbibliotek: verksamhet och visioner: betänkande, Statens offentliga utredningar, 2003:129*. Stockholm: Fritzes offentliga publikationer, 2003.
- Keats-Rohan, K. S. B., and Linacre College (University of Oxford). Unit for Prosopographical Research. *Resourcing sources*. Oxford: Unit for Prosopographical Research, 2002.
- Knorr-Cetina, Karin D. *Epistemic cultures: how the sciences make knowledge*. Cambridge, Mass.: Harvard University Press, 1999.
- Kronman, Ulf, and John Parnefjord. "Resource Description Framework. Metadata för Internet." *Human IT*, no. 4 (1999).
- "Kulturellt innehåll - digitalisering, elektronisk tillgång och digitalt bevarande." edited by Europeiska gemenskapernas kommission (EU-kommissionen): Europeiska unionens officiella tidning, 31 aug. 2006.
- Lancashire, Ian. "Early Books, RET Encoding Guidelines, and the Trouble with SGML." Paper presented at the Scriptorium conference, Calgary, November 1995.
- . "Using TACT with electronic texts: a guide to text-analysis computing tools." New York: Modern Language Association of America, 1996.
- Langerth Zetterman, Monica. *Kvinnors mötesplatser och sociala och kulturella tillgångar kring sekelskiftet 1900. Beskrivning av prosopografiskt material inom projektet Formering för offentlighet*. 1 ed. Vol. 41, *Rapporter från Forskningsgruppen för utbildnings- och kultursociologi*. Uppsala: Forskningsgruppen för utbildnings- och kultursociologi, Under utgivning.
- . "Märkvärt. Noteringar om en prosopografi över kvinnor i offentlighet kring sekelskiftet 1900." In *Fältanteckningar: utbildnings- och kultursociologiska texter tillägnade Donald Broady*, edited by Mikael Börjesson, Ingrid Heyman, Monica Langerth Zetterman, Esbjörn Larsson, Ida Lidegran and Mikael Palme. Uppsala: Forskningsgruppen för utbildnings- och kultursociologi ILU Uppsala universitet, 2006.
- . "Using Markup for Multivariate Analyses in the Prosopographical Study "Formation for the Public Sphere"." Paper presented at the ACH/ALLC Conference, Victoria University, Canada, June 15 - June 18 2005.
- Larsson, Esbjörn. *Att studera titlar som kapital: en teoretisk diskussion kring militära titlar i 1700-talets svenska samhälle*, 2003.
- . *Från adlig uppfostran till borgerlig utbildning: Kungl. krigsakademien mellan åren 1792 och 1866*, *Studia historica Upsaliensia*, 220. Uppsala: Acta Universitatis Upsaliensis: Uppsala University Library distributör, 2005.
- Levy, David M. *Scrolling forward: making sense of documents in the digital age*. 1. ed. New York: Arcade Pub, 2001.
- Linné, Agneta. "Lärarinnor, pedagogiskt arbete och offentlighet. Kvinnliga pionjärs liv och strategier kring år 1900 - ett jämförande perspektiv."

- Paper presented at the Svenska Historikermötet, Uppsala, 22-24 april 2005.
- Linné, Agneta, and Kerstin Skog-Östlin, eds. *Biografi och kollektivbiografi som historisk metod: NFPF:s kongress 15-18 mars 2001 i Stockholm, Meddelanden från Forum för pedagogisk historia, 5*. Uppsala: Forum för pedagogisk historia SEC ILU Univ., 2001.
- Lopez, Silbano de. "Kungliga Högre lärarinneseminariet och flickskolans framväxt." In *Sekelskiftets utmaningar: essäer om välfärd, utbildning och nationell identitet vid sekelskiftet 1900*, edited by Ann-Katrin Hatje, 181-98. Stockholm: Carlsson, 2002.
- Lundström, Nils Styrbjörn. *Svenska kvinnor i offentlig verksamhet: porträtt och biografier*. Uppsala, 1924.
- Marshall, C. C, and Bernheim B. A. J. "Exploring the Relationship between Personal and Public Annotations." Paper presented at the ACM/IEEE Joint Conference on Digital Libraries (JCDL04), Tucson, Arizona, June 7–11 2004.
- Marshall, Catherine, and Frank Shipman. "Which Semantic Web?" Paper presented at the Fourteenth ACM Conference on Hypertext and Hypermedia, Nottingham UK 2003.
- McCarty, Willard. *Humanities computing*. Basingstoke; New York: Palgrave Macmillan, 2005.
- McDermott, Drew, and Dejing Dou. "Representing Disjunction and Quantifiers in RDF." Paper presented at the First International Semantic Web Conference, June 9-12 2002.
- McGann, Jerome J. "Making texts of many dimensions." In *A companion to digital humanities*, edited by Susan Schreibman, Raymond George Siemens and John Unsworth. Malden, MA; Oxford: Blackwell Pub., 2004.
- . *Radiant textuality: literary studies after the World Wide Web*. New York; Basingstoke: Palgrave, 2001.
- McKenzie, D. F. *Bibliography and the sociology of texts*. Cambridge: Cambridge University Press, 1999.
- Melkersson, Hans. "Innehållsdesign. Om principer för märkning och lagring av innehåll." In *Fältanteckningar: utbildnings- och kultursociologiska texter tillägnade Donald Broady*, edited by Mikael Börjesson, Ingrid Heyman, Monica Langerth Zetterman, Esbjörn Larsson, Ida Lidegran and Mikael Palme. Uppsala: Forskningsgruppen för utbildnings- och kultursociologi ILU Uppsala universitet, 2006.
- Morrison, Alan, Michael Popham, and Karen Wikander. *Creating and Documenting Electronic Texts: A Guide to Good Practice*. Oxford: Oxford Text Archive, 2000.
- Müller, Eva, Uwe Klosa, Stefan Andersson, and Peter Hansson. "The DiVA Project - Development of an Electronic Publishing System." *D-Lib Magazine* 9, no. 11 (2003).
- Mylonas, Ellen, and Allen H. Renear. "The Text Encoding Initiative at 10: Not Just an Interchange Format Anymore – But a New Research Community." *Computers and the Humanities* 33, no. 1-2 (1997): 1-9.

- Ong, Walter J., and Lars Fyhr. *Muntlig och skriftlig kultur: teknologiseringen av ordet*. 2:a uppl. ed. Göteborg: Anthropos, 1991.
- Pepper, Steve. "The TAO of topic maps." 2002.
- Pettersson, Karin. *Den manliga heterosexuella normen: en analys av SAB-systemet utifrån ett feministiskt och queerteoretiskt perspektiv*, Magisteruppsats i biblioteks- och informationsvetenskap vid Bibliotekshögskolan/Biblioteks- och informationsvetenskap, 2001:73. Borås: Högsk. i Borås Bibliotekshögskolan/Biblioteks- och informationsvetenskap, 2001.
- Pickering, Andrew. *Science as practice and culture*. Chicago: Univ. of Chicago Press, 1992.
- Piez, Wendell. "Beyond the "descriptive vs procedural" distinction." Paper presented at the Extreme Markup Languages 2001 2001.
- Rahtz, Sebastian, Norman Walsh, and Lou Burnard. "A unified model for text markup: TEI, Docbook, and beyond." Paper presented at the XML Europe 2004, Amsterdam, The Netherlands, 18-21 April 2004.
- Raymond, Eric S. *The cathedral and the bazaar. Musings on Linux and open source by an accidental revolutionary*. Rev. ed. Sebastopol, Calif.: O'Reilly, 2001.
- Renear, Allen H. "The Descriptive/Procedural Distinction is Flawed." *Markup Languages: Theory and Practice* 2, no. 4 (2000): 411-20.
- . "Out of Praxis: Three (Meta)Theories of Textuality." In *Electronic text: investigations in method and theory*, edited by Kathryn Sutherland, 107-26. Oxford: Clarendon, 1997.
- . "Text Encoding." In *A companion to digital humanities*, edited by Susan Schreibman, Raymond George Siemens and John Unsworth, 218-39. Malden, MA; Oxford: Blackwell Pub., 2004.
- . "Text from Several Different Perspectives, the Role of Context in Markup Semantics." Paper presented at the Proceedings of the 2003 Conference on Computers, Literature, and Philology, Florence: University of Florence 2003.
- Renear, Allen H., David G. Durand, and Ellen Mylonas. "Refining our notion of what text really is." In *Research in Humanities Computing 4: Selected papers from the 1992 ALLC/ACH Conference.*, edited by Susan M. Hockey and Nancy Ide, 263-80. Oxford: Oxford University Press, 1997.
- Roos, J.P., and Anna Rotkirch. "Fält i skuggan av fält: sensocialismens dubbel-liv." In *Kulturens fält: Cultural fields: en antologi*, 453-66. Göteborg: Daidalos, 1998.
- Rönnbäck, Josefin. *Politikens genusgränser: den kvinnliga rösträttsrörelsen och kampen för kvinnors politiska medborgarskap 1902-1921*, Atlas akademi. Stockholm: Atlas, 2004.
- Rönholm, Tord. *Kunskapens kvinnor: sekelskiftets studentkor i mötet med den manliga universitetsvärlden*, Forskningsrapporter från Institutionen för historiska studier vid Umeå universitet, 12. Umeå: Institutionen för historiska studier Univ. distributör, 1999.
- Sapiro, Gisele. "The structure of the French literary field during the German Occupation (1940–1944): a multiple correspondence analysis." *Poetics* 30 (2002): 387-402.

- Schreibman, Susan, Raymond George Siemens, and John Unsworth. *A companion to digital humanities*, *Blackwell companions to literature and culture*; 26. Malden, MA; Oxford: Blackwell Pub., 2004.
- Shadbolt, Nigel, Wendy Hall, and Tim Berners-Lee. "The Semantic Web Revisited." *IEEE Intelligent Systems* 21, no. 3 (2006): 96-101.
- Shillingsburg, Peter. "Principles for Electronic Archives, Scholarly Editions, and Tutorials." In *The literary text in the digital age*, edited by Richard J. Finneran. Ann Arbor: University of Michigan Press, 1996.
- Shillingsburg, Peter L. *From Gutenberg to google: electronic representations of literary texts*. Cambridge: Cambridge Univ. Press, 2006.
- Skog-Östlin, Kerstin. *Att bryta ny mark kvinnors bruk av läroverkslärarutbildning omkring 1900, Rapporter från Pedagogiska institutionen, Örebro universitet, 11*. Örebro: Örebro universitet Pedagogiska institutionen, 2005.
- Sperberg-McQueen, C. M. "Classification and its Structures." In *A companion to digital humanities*, edited by Susan Schreibman, Raymond George Siemens and John Unsworth, 160-76: Blackwell Pub., 2004.
- . "Rabbit/duck grammars: a validation method for overlapping structures." Paper presented at the Extreme Markup Languages 2006, Montréal, Québec 2006.
- . "Representation of overlapping structures." Paper presented at the Extreme Markup Languages 2007, Montréal, Québec 2007.
- Sperberg-McQueen, C. M., and L. Burnard. "The Design of the TEI Encoding Scheme." In *Text encoding initiative: background and context*, edited by Nancy Ide and Jean Vâeronis, 17-40. Dordrecht; London: Kluwer Academic, 1995.
- Statskontoret. "Öppen programvara." In *Rapport 2003:8*. Stockholm: Statskontoret, 2003.
- . "Öppen programvara. Erfarenheter av produkter som bygger på öppen källkod inom förvaltningen." In *Rapport 2004:21*. Stockholm: Statskontoret, 2004.
- Stone, Lawrence. "Prosopography." In *Historical Studies Today*, edited by Felix Gilbert and Stephen R. Graubard, 107-40. New York: Norton, 1972.
- Svedjedal, Johan. *Den sista boken*. Stockholm: Wahlström & Widstrand, 2001.
- . "En svensk Litteraturbank? Utredning för Svenska Akademien, avlämnad i januari 2003." 2003.
- Säljö, Roger, and Maria Södling. "Utbildning på vetenskaplig grund: röster från fältet." In *Högskoleverkets rapportserie, 2006:46 R*, 64. Stockholm: Högskoleverket, 2006.
- Tennison, Jeni, and Wendell Piez. "The Layered Markup and Annotation Language (LMNL)." Paper presented at the Extreme Markup Languages 2002, Montréal, Québec 2002.
- Thompson, John B. *Books in the digital age: the transformation of academic and higher education publishing in Britain and the United States*. Cambridge: Polity, 2005.
- Thorsell, Lennart. "Den svenska parnassens "demokratisering" och de folkliga bildningsvägarna." In *Samlaren*, 53-135: Uppsala, 1958.

- Tosh, John. *Historisk teori och metod*. 2., rev. uppl. ed. Lund: Studentlitteratur, 2000.
- Ullman, Annika. *Stiftarinnegenerationen: Sofi Almqvist, Anna Sandström, Anna Ahlström*. Stockholm: Stockholmia, 2004.
- Vanhoutte, Edward. "An introduction to the TEI and the TEI Consortium." *Literary and Linguistic Computing* 19, no. 1 (2004): 9-19.
- Werner, Karl Ferdinand. "'L'apport de la prosopographie à l'histoire sociale des élites'". In *Family trees and the roots of politics: the prosopography of Britain and France from the tenth to the twelfth century*, edited by K. S. B. Keats-Rohan, 1-21. Woodbridge: Boydell Press, 1997.
- Vetenskapsrådet. "Om forskningens infrastrukturer inom humaniora och samhällsvetenskap i Sverige." Vetenskapsrådet, 2005.
- Wieselgren, Greta. *Den höga tröskeln: kampen för kvinnas rätt till ämbete, Kvinnohistoriskt arkiv*, 7. Lund: Gleerup, 1969.
- Winograd, Terry. "Thinking machines: Can there be? Are we?" In *The Boundaries of Humanity: Humans, Animals, Machines*, edited by James Sheehan and Morton Sosna, 198-223. Berkeley: University of California Press, 1991.
- Österberg, Eva, and Christina Carlsson Wetterberg. *Rummet vidgas: kvinnor på väg ut i offentligheten 1880-1940*. Stockholm: Atlantis, 2002.