

MAGISTERUPPSATS I BIBLIOTEKS- OCH INFORMATIONSVETENSKAP
VID BIBLIOTEKS- OCH INFORMATIONSVETENSKAP/BIBLIOTEKSHÖGSKOLAN
2003:56

Open Archives Initiative - om ”öppna arkiv” och den sociala kontextens betydelse

Lillemor Åkerman

© **Författaren/Författarna**

Mångfaldigande och spridande av innehållet i denna uppsats
– helt eller delvis – är förbjudet utan medgivande av författaren/författarna.

Svensk titel:	Open Archives Initiative – om ”öppna arkiv” och den sociala kontextens betydelse
Engelsk titel:	Open Archives Initiative – about ”open archives” and the influence of the social context
Författare:	Lillemor Åkerman
Kollegium:	Kollegium 2
Färdigställt:	VT 2003
Handledare:	Jan Buse. Kollegium 2

Abstract: There is an ongoing evolution in scholarly communication. The traditional way of communicating scholarly results through scientific journals is beginning to be replaced by other means of dissemination. One new channel for scholars to communicate their results is digital publication in so-called open archives. An open archive is a kind of full text database, where a scholar may archive/publish his or her articles, as a complement or a substitute to publishing in scientific journals. There are discipline-oriented archives and institutional archives. The Open Archives Initiative (OAI), set up in 1999, is a major development aimed at heightening interoperability between different open archives.

The aim of this thesis is twofold. The first aim is to give an overview of the development of the OAI, mainly from a knowledge organizational perspective. The second purpose is to examine open archives from a historical, social and cultural perspective. The open archives movement has not been the immediate success that many of the advocators had hoped. It is also said that scholars in the humanities and the social sciences are more reluctant to self-archiving than scholars in the natural sciences. This thesis try to find out more about why scholars are reluctant and to examine whether it is true that scholars in the humanities and the social sciences are more reluctant, and if it's true, why this is so. This study compiles some weaknesses with the OAI retrieval model that have been observed lately. Many of these weaknesses are related to human aspects of the retrieval model, for example issues related to the creation of metadata.

The results show that scholars in the humanities and the social sciences have been slow to accept open archives, as a reliable channel of publication and dissemination and that the reasons for this may be historical, social and cultural. The study compiles different barriers that may prevent or delay a more widespread usage of open archives. Some barriers may affect scholars in all disciplines; other may mostly affect scholars in the humanities and the social sciences.

This study proposes that historical, social and cultural issues have to be taken into consideration when new information retrieval systems are developed.

Nyckelord: Open Archives Initiative, vetenskaplig kommunikation, vetenskaplig publicering, humaniora, samhällsvetenskap, social kontext, metadata, egenarkivering

”The way forward is paradoxically to look not ahead, but to look around.”

(Brown & Duguid 2000, s. 8)

Innehållsförteckning

1. Inledning	s. 6
1.1. Ämnesval.....	s. 6
1.2. Bakgrund och problembeskrivning.....	s. 6
1.3. Syfte och frågeställningar	s. 8
1.4. Begränsningar	s. 9
1.5. Disposition.....	s. 10
2. Teori	s. 10
2.1. Definitioner.....	s. 10
2.1.1. Kunskapsorganisation	s. 10
2.1.2. Vetenskaplig kommunikation.....	s. 11
2.1.3. Digitala bibliotek.....	s. 12
2.2. Teoretisk referensram.....	s. 14
2.3. Kunskapsöversikt.....	s. 18
2.3.1. OAI-studier.....	s. 18
2.3.2. Kunskapsorganisation	s. 18
2.3.3. Vetenskaplig kommunikation.....	s. 19
2.3.4. Digitala bibliotek	s. 20
3. Metod	s. 21
3.1. Metodval.....	s. 21
3.2. Informationssökning	s. 22
4. Open Archives Initiative	s. 23
4.1. Historisk bakgrund.....	s. 23
4.2. Alternativ publicering – olika aktörer.....	s. 25
4.2.1. <i>Scholarly Publishing and Academic Resources Coalition (SPARC)</i>	s. 25
4.2.2. <i>Public Library of Science (PloS)</i>	s. 25
4.2.3. <i>Budapest Open Access Initiative (BOAI)</i>	s. 26
4.2.4. <i>FIGARO</i>	s. 26
4.3. E-printarkiv – om bakgrunden till OAI.....	s. 26
4.4. Open Archives Initiative	s. 27
4.5. OAI Protocol for Metadata Harvesting 2.0.....	s. 29
4.5.1. <i>Metadata</i>	s. 29
4.5.2. <i>Interoperabilitet</i>	s. 30
4.5.3. <i>Metadata enligt OAI: s modell</i>	s. 33
4.5.4. <i>Metadatapost enligt OAI: s modell</i>	s. 34
4.5.5. <i>Protokollfrågor</i>	s. 36
4.5.6. <i>Kritiska röster</i>	s. 36
4.6. Sammanfattning avsnitt 4	s. 39
5. Vägen mot ”öppna arkiv”	s. 41
5.1. En ny modell, men för vem?.....	s. 41
5.2. Den historiska och sociala kontextens betydelse	s. 44
5.3. Om attityder till och tänkbara barriärer mot egenarkivering	s. 47
5.3.1. <i>Bevarandeaspekter</i>	s. 48
5.3.2. <i>Kvalitetskontroll</i>	s. 48
5.3.3. <i>Upphovsrätt</i>	s. 50

5.3.4. Val av publiceringsform	s. 54
5.3.5. Tekniska faktorer	s. 56
5.3.6. Ekonomiska faktorer	s. 58
5.4. Sammanfattning avsnitt 5	s. 59
6. Diskussion.....	s. 60
6.1. Slutsatser	s. 66
7. Sammanfattning	s. 67
8. Litteraturlista	s. 71

1. Inledning

1.1. Ämnesval

Alla som någon gång har skrivit en magisteruppsats vet att det är en uppgift som tar tid och som kräver en hel del tankeverksamhet. Detta gäller inte minst i inledningsskedet när val av ämne ska ske. Utgångspunkten för mig var att hitta ett ämne med anknytning till elektronisk publicering, som samtidigt kunde anses höra hemma inom det kollegium jag tillhör - Kunskapsorganisation. Anledningen till att jag ville skriva om elektronisk publicering var att jag parallellt med mina studier i Biblioteks- och informationsvetenskap har läst en givande kurs kallad Elektronisk publicering i praktik och teori. Min första tanke var att skriva om katalogisering av e-böcker. Under det att jag läste in mig på litteratur i detta ämne kom jag att läsa en text av Mats Cavallin (2002), chef för Göteborgs universitets Digitala bibliotek. I texten *e-bokens lov* skriver Mats Cavallin om övergången från tryckta böcker till e-böcker och om den digitalisering av vetenskaplig litteratur som pågår sedan ett antal år tillbaka. Denna text fick mig att ändra inriktning och i stället börja utforska ämnet "elektronisk publicering av vetenskaplig litteratur". Ganska omgående dök i flera sammanhang olika förkortningar upp: OAI, SPARC, BOAI, PLoS, vilka alla visade sig vara olika slags aktörer i en tämligen omfattande omstrukturering av den vetenskapliga kommunikationen, i vilken elektronisk publicering är en av grundbultarna. Denna omstrukturering kallas ibland för ett pågående paradigmskifte inom vetenskaplig kommunikation, vilket ytterligare spädde på min nyfikenhet. Jag lockades på detta sätt in det som kom att bli ämnet för min magisteruppsats – Open Archives Initiative.

1.2. Bakgrund och problembeskrivning

Digitala bibliotek har nu funnits med oss ett antal år. Ett digitalt bibliotek är ingalunda någon statiskt företeelse, utan dess form och innehåll förändras kontinuerligt. De digitala biblioteken ställs oavbrutet inför nya utmaningar och en av flera nya uppgifter är att vara delaktiga i uppbyggnaden av det som på engelska kallas "open archives", i denna uppsats benämnda "öppna arkiv".

The Open Archives Initiative, hädanefter kallat OAI, bildades i USA 1999. Bakgrunden till initiativet var ett ökat intresse för alternativa publiceringsformer inom vetenskaplig kommunikation. Målet med OAI:s arbete är att forskare ska arkivera/publicera sina resultat i "öppna arkiv", en slags fulltextdatabaser, som ersättning för eller komplement till publicering i kommersiella vetenskapliga tidskrifter. Ett "öppet arkiv" kan vara en forskares personliga hemsida, där han eller hon publicerar allt eget material, en ämnesdatabas för vetenskapligt material inom en viss disciplin eller en databas för allt vetenskapligt material som skapas vid ett universitet, ett s.k. institutionsarkiv. Dessa arkiv ska sedan, enligt de största optimisterna, knytas samman i ett enda stort vetenskapligt världsarkiv. Man skulle kunna se det som en slags frihetsrörelse där målet är att forskare ska återta makten över forskningsresultaten, i stället för att som nu skänka bort dem till kommersiella förlag. Steven Harnad, en av de "öppna arkivens" främsta förespråkare, har uttryckt saken på följande sätt:

As soon as all refereed journal articles are self-archived by their authors in their institution's eprint archive, the literature is freed from all access barriers and impact barriers. Self-archiving could be done virtually overnight. The day after, all refereed research becomes freely accessible online to researchers the world over. (Harnad 2001)

Vissa förespråkare tycks i mina ögon ha drabbats av en viss fartblindhet och det finns då inga gränser för vad dessa alternativa publiceringsformer kan tänkas leda till. Så kan man t.ex. på Budapest Open Access Initiatives hemsida läsa följande ord:

Removing access barriers to this literature [peer-reviewed journal literature; min anm.] will accelerate research, enrich education, share the learning of the rich with the poor and the poor with the rich, make this literature as useful as it can be, and lay the foundation for uniting humanity in a common intellectual conversation and quest for knowledge. (Budapest Open Access Initiative 2002)

De finns de som förutspår eller hoppas att "öppna arkiv" och andra alternativa publiceringsformer ska bli det dominerande sättet för forskare att publicera sig på. Samsökning i de "öppna arkiven" sker, enligt den modell som OAI utarbetat, med hjälp av ett protokoll kallat *the OAI Protocol for Metadata Harvesting*. Protokollet bygger på traditionella kunskapsorganisatoriska begrepp som lagring, sökning och återvinning av dokument, i det här fallet digitala dokument. Metadata (kortfattade beskrivningar av dokumenten) enligt ett standardiserat format, är det kitt som håller ihop alla olika "öppna arkiv". OAI: s arbete med *the OAI Protocol for Metadata Harvesting* gick från tanke till handling på mycket kort tid. Arbetet initierades i oktober 1999 och 14 månader senare släpptes den första versionen. Protokollet som sådant tycks fungera, även om det med all säkerhet kommer att dyka upp en och annan "barnsjukdom".

Något som förvånar, och till viss del bekymrar många av förespråkarna, är det motstånd mot egenpublicering som verkar finnas hos en del forskare. Övergången till en ny modell för vetenskaplig kommunikation har enligt vissa debattörer varit långsammare och mindre omfattande än både väntat och önskat. Ett uttalande av Steven Harnad får illustrera det motstånd som tycks finnas inom vissa discipliner:

But although these physicists have shown the way to free the refereed research literature, authors in other disciplines have been slow to realize that the system can work for them to. (Harnad 2001)

Trots att det har skapats en del "öppna arkiv" även inom humaniora och samhällsvetenskap, antyder en del debattörer att forskare inom just dessa områden verkar ha en förhållandevis avvaktande inställning till elektronisk egenpublicering. Eftersom jag har mina rötter inom humaniora och samhällsvetenskap började jag fundera över om forskare inom dessa discipliner verkligen förhåller sig mer avvaktande och vad det i så fall kan bero på.

Även för universitets- och högskolebibliotek sägs "öppna arkiv" vara ett fördelaktigt alternativ till traditionell publicering. Bibliotekens kostnader för prenumerationer på vetenskapliga tidskrifter har ökat långt utöver många andra prisstegringar. Det skulle innebära stora ekonomiska vinster för universitetsbiblioteken om dyra prenumerationer

skulle kunna bytas ut mot förhållandevis billigare arkiv. Det är dock inte bara en ekonomisk fråga. "Öppna arkiv" sägs leda till ökad tillgänglighet, inte bara för forskare utan även för folk i allmänhet. Dessutom menar en del att ett institutionsarkiv på ett påtagligt sätt kan visa vilken forskning som bedrivs vid ett universitet. I stället för att sprida ett universitets vetenskapliga resultat på ett stort antal vetenskapliga tidskrifter, kan alla resultat samlas på ett och samma ställe. Institutionsarkiv skulle därmed kunna fungera som en slags kvalitetsindikator och bidra till att synliggöra och höja universitetens status (Crow 2002; Johnson 2002).

Omfattande egenarkivering skulle innebära en omstrukturering av den traditionella vetenskapliga publiceringskedjan, författare - kommersiella förlag/lärda sällskap - bibliotek. När huvuddelen av ett biblioteks bestånd av vetenskapliga tidskrifter är förlagsägda, är bibliotekets huvudsakliga uppgifter att välja ut och i fysisk eller elektronisk form tillhandahålla materialet. Visserligen har många universitet redan idag en viss förlagsverksamhet, som utgivare av exempelvis litteratur inom humaniora och samhällsvetenskap. Genom en utveckling av "öppna arkiv" skulle förlagsrollen förstärkas och universitets- och högskolebiblioteken få ett större ansvar för produktion, lagring och bevarande av elektroniska dokument. En sådan utveckling innebär onekligen nya utmaningar för dem som arbetar vid de digitala biblioteken.

Det pågår världen över en intensiv debatt om den obalans som anses råda inom det traditionella publiceringsparadigmet, med oacceptabla prenumerationskostnader och begränsad access till forskningsresultat som följd. Även i Sverige har man under senare år börjat diskutera frågan och det finns en hel del skrivet på svenska om krisen inom vetenskaplig publicering. När det gäller OAI och andra försök att skapa tekniker och organisatoriska förutsättningar för en utveckling av "öppna arkiv" finns det, så vitt jag har kunnat bedöma, mycket lite skrivet på svenska. Vid flera svenska universitet och högskolor pågår dock försök med att skapa "öppna arkiv". Jag anser att "öppna arkiv" är en mycket intressant företeelse, som ligger i tiden och som utifrån ett flertal olika aspekter förtjänar att uppmärksammas och diskuteras.

1.3. Syfte och frågeställningar

Jag har två syften med denna uppsats. Det första syftet är att utifrån ett kunskapsorganisatoriskt perspektiv undersöka hur "öppna arkiv" fungerar utifrån aspekter som lagring, sökning och återvinning av information. Med tanke på vad arkiven ska innehålla, nämligen delar av en vetenskaplig kommunikation som till stor del är en social process, kommer jag även att försöka finna förklaringar till eventuella skillnader mellan olika discipliners inställning till "öppna arkiv". Man bör, enligt min åsikt, inte stirra sig blind på standarder och tekniska lösningar, utan även väga in andra faktorer när man försöker skapa nya publiceringsformer. Det andra syftet med uppsatsen blir därför att undersöka "öppna arkiv" utifrån ett socialt, historiskt och kulturellt perspektiv. I anslutning till båda mina syften ämnar jag även undersöka de digitala bibliotekens roll/roller i uppbyggnaden av "öppna arkiv".

Att sociala, historiska och kulturella aspekter kan ha betydelse även inom kunskapsorganisation har bland andra Christine L. Borgman och Jack Andersen argumenterat för. Borgman (2000a) belyser relationen mellan digitala bibliotek och vetenskaplig kommunikation. Borgman nämner i sin artikel inte uttryckligen "öppna arkiv", men gör reflektioner kring digitala bibliotek och vetenskaplig kommunikation

som kan ha relevans även för denna uppsats syfte. Hon menar att man genom att studera enligt vilka kriterier forskare väljer publiceringsform, kan få kunskap som kan användas vid de digitala bibliotekens utveckling av alternativa publiceringsformer för vetenskapliga arbeten.

Andersen (2002) argumenterar för en breddning av den gängse definitionen av begreppet kunskapsorganisation. Kunskapsorganisation har traditionellt definierats som ett arbete med att skapa tekniker, regler och standarder för dokumentbeskrivningar och organisation av dessa beskrivningar, med syftet att underlätta återvinning av de beskrivna dokumenten. Andersen menar att man måste kombinera studier av kunskapsorganisatoriska problem med analyser av samhällets kunskapsproduktion. Eftersom vetenskaplig kommunikation är en viktig del av samhällets kunskapsproduktion, finns det utifrån ovanstående perspektiv, enligt min mening skäl att även undersöka olika aspekter av vetenskaplig kommunikation vid kunskapsorganisatoriskt inriktade studier av ”öppna arkiv”.

De frågeställningar jag vill söka svar på i denna uppsats är följande:

- Hur fungerar ”öppna arkiv” utifrån ett kunskapsorganisatoriskt perspektiv, d.v.s. vilka metoder används vid lagring, sökning och återvinning av information?
- Vilka eventuella barriärer finns det mot en vidareutveckling av ”öppna arkiv” och andra alternativa publiceringsformer?
- Om forskare inom humaniora och samhällsvetenskap i större utsträckning förhåller sig avvaktande till ”öppna arkiv” än forskare inom naturvetenskap, vilka tänkbara orsaker kan det i så fall finnas till att det förhåller sig så?

1.4. Begränsningar

Det finns tre huvudsakliga aktörer i den vetenskapliga kommunikationskedjan - forskare, förlag och bibliotek. Alla dessa tre aktörers agerande har stor betydelse för vilken riktning den vetenskapliga kommunikationen kommer att ta framöver. Utöver kunskapsorganisatoriska aspekter kommer jag dock huvudsakligen att undersöka ”öppna arkiv” utifrån forskarnas och i viss mån även bibliotekens agerande. Jag kommer att beröra förlagens roll, men i begränsad omfattning. Jag är medveten om att det inte går att ge en heltäckande bild av det ”ställningskrig” som nu pågår, med forskare och bibliotek på ena sidan barriärerna och förlagen på den andra, utan att redogöra för alla parter agerande. Jag har dock, av utrymmesskäl, men även utifrån andra ställningstaganden, valt att koncentrera mig på dem som ska *använda* arkiven, eftersom jag tror att det är dessa personer som i slutändan kommer att avgöra de ”öppna arkivens” framtid.

Uppsatsen kan därmed anses kretsa kring tre begrepp - kunskapsorganisation, vetenskaplig kommunikation och digitala bibliotek. Dessa tre begrepp kommer jag att definiera i teoriavsnittet.

1.5. Disposition

Uppsatsen är i kommande text disponerad enligt följande. Först kommer ett teoriavsnitt där jag definierar tre grundläggande begrepp, sätter upp en teoretisk referensram för uppsatsen samt ger en mindre översikt över kunskapsområden av intresse för min uppsats. Därefter följer ett metodavsnitt, där jag redogör för val av metod samt hur jag har gått tillväga för att finna information. På det följer två mer omfattande avsnitt. Det första handlar om bakgrunden till och framväxten av Open Archives Initiative och om den återvinningsmodell som ”öppna arkiv” bygger på. Det andra har jag valt att kalla *Vägen mot ”öppna arkiv”*. I det avsnittet diskuterar jag vetenskaplig kommunikation och olika barriärer som kan tänkas försvåra en övergång till alternativa publiceringsmodeller. Därefter följer ett diskussionsavsnitt som avslutas med en sammanställning av de slutsatser jag har dragit av mitt arbete. Avslutningsvis finns ett sammanfattande avsnitt för hela uppsatsen.

2. Teori

I följande avsnitt kommer jag att definiera tre, för min uppsats, viktiga begrepp - kunskapsorganisation, vetenskaplig kommunikation och digitala bibliotek.

2.1. Definitioner

2.1.1. Kunskapsorganisation

Kunskapsorganisation är ett begrepp med mycket gamla anor. Alltsedan människan började teckna ner sina tankar har det funnits ett behov av att på ett systematiskt vis ”märka upp” och organisera denna nedtecknade kunskap, för att göra den möjlig att återfinna. Bibliotekskatalogen är ett av de mest tydliga exemplen på systematiskt ordnad kunskap. Genom århundraden har människan utvecklat allt mer raffinerade metoder för att beskriva och återfinna olika slags information och kunskap. Genom utvecklingen av datorer, Internet och webben sker detta arbete idag på ett radikalt annorlunda sätt än tidigare. Men även om tekniken har förändrats är syftet med kunskapsorganisation det samma, att skapa ordning i en allt snabbare växande informationsmängd. James Anderson definierar kunskapsorganisation på följande sätt:

The description of documents, their content, features and purpose, and the organization of their descriptions so as to make these documents and their parts accessible to persons seeking them or the messages that they contain. Knowledge organization encompasses every type and method of indexing, abstracting, cataloguing, classification, records management, bibliography and the creation of textual or bibliographic databases for information retrieval... In the context of library and information science, the organization of knowledge (often called the organization of information) is the organization of documented messages in which knowledge or information is represented.
(Anderson 1996, s. 336)

Detta får nog anses vara den rådande definitionen. Utbildningen i Biblioteks- och informationsvetenskap vid Högskolan i Borås är uppdelad i fyra kollegier, varav ett

kallas Kunskapsorganisation. De delmoment som är specifika för Kunskapsorganisation, i förhållande till de andra tre kollegierna, kallas dokumentrepresentation, ämnesrepresentation och information retrieval och speglar i stor utsträckning den syn på kunskapsorganisation som kommer till uttryck i James Andersons definition. I grundutbildningen har man som jag ser det, trots en gemensam första termin för de fyra kollegierna, valt att låta undervisningen i kunskapsorganisation stå relativt isolerad från andra aspekter som i mitt tycke kan ha relevans för kunskapsorganisatoriska frågeställningar. Även om James Andersons definition av kunskapsorganisation troligtvis är den mest accepterade, finns det forskare som anser att kunskapsorganisation bör sättas in i ett större sammanhang. En av dem är Jack Andersen (2002). Jag kommer att redogöra för Andersens tankar kring en bredare definition av kunskapsorganisation i avsnittet Teoretisk referensram.

2.1.2. Vetenskaplig kommunikation

Begreppet vetenskaplig kommunikation myntades så sent som i mitten av 1970-talet. Likaså var det först vid den här tiden som man började diskutera vetenskaplig kommunikation som system (Lindholm-Romantschuk 1998, s. 21). Men även om det inte kallades för vetenskaplig kommunikation tidigare, har forskare sedan lång tid tillbaka utbytt tankar och idéer med varandra. Hur viktig kommunikationen är för forskningens framsteg har Meadows uttryckt med följande ord:

Communication lies at the heart of research. It is as vital for research as the actual investigation itself, for research cannot properly claim that name until that it has been scrutinized and accepted by colleagues. This necessarily requires that it be communicated. (Meadows 1998, s. ix)

Forskare kommunicerar både genom att utbyta erfarenheter och resultat med varandra under arbetets gång och genom att publicera sina slutresultat. Vetenskaplig kommunikation kan därmed vara både informell och formell. Informell kommunikation definieras som allt utbyte forskare emellan "som inte offentliggörs genom en formell publikationsprocess" (Lindholm-Romantschuk 1998, s. 10; *min översättning*). Samtal, telefonsamtal, privat korrespondens, besök på andra universitet och konferenser är alla exempel på informell kommunikation (Lindholm-Romantschuk 1998; Rowland 1997). Ett annat uttryck för den aktivitet som sker genom informell kommunikation är "invisible colleges" (Crane, se Rowland 1997).

Formell kommunikation å andra sidan är lika med "all kommunikation som är kodifierad och offentliggjord genom en formell publikationsprocess" (Lindholm-Romantschuk 1998, s. 10; *min översättning*). Traditionellt sett har den formella kommunikationen utgjorts av tryckt material - tidskrifter, monografier, konferensrapporter o.s.v. Vissa väljer att kalla den formella kommunikationen för vetenskaplig publicering, som då ses som slutprodukten av vetenskaplig kommunikation i stort (se t.ex. Halliday 2001). Jag kommer att försöka skilja mellan arkivering och publicering. Det kommer dock inte att bli lätt, då de båda orden ofta används synonymt. Med arkivering kommer jag att avse forskare som tillgängliggör sitt vetenskapliga material i nätverksbaserade "öppna arkiv". Med publicering kommer jag att avse ett offentliggörande av vetenskapliga resultat i traditionella tryckta eller elektroniska tidskrifter, ägda av förlag eller lärda sällskap. I de fall det är tveksamt vad

som menas kommer jag att skriva arkivering/publicering. Denna åtskillnad har jag valt att göra p.g.a. att det fortfarande är en kontroversiell fråga huruvida arkivering ska jämföras med publicering. Rowland (1997; *min översättning*) hävdar t.ex. att det är viktigt att ”skilja mellan en akademisk debatt och vetenskaplig publicering”. En del av den akademiska debatten är, enligt Rowland, den allt mer utbredda diskussionen forskare emellan som Internet gett möjlighet till. Vetenskaplig publicering däremot, innefattar en rigorös kvalitetskontroll och ska syfta till ett långsiktigt bevarande av tillförlitliga forskningsresultat (Rowland 1997). Halliday (2001) har, genom en enkätundersökning bland välutbildade personer, försökt skapa en definition av vetenskaplig publicering och finna kriterier för när ett dokument kan anses *publicerat* i vetenskaplig bemärkelse. Enligt Rowlands och Hallidays sätt att se går det en klar gräns mellan informell och formell kommunikation. Andra (t.ex. Lindholm-Romantschuk 1998, s. 10) anser att gränsen mellan informell och formell kommunikation aldrig har varit knivskarp och att den ökade användningen av elektroniska medier ytterligare kommer att suddas ut gränsen, genom att dessa används för både formell och informell kommunikation. Steven Harnad menar t.o.m. (enligt Lindholm-Romantschuk 1998, s.10) att distinktionen mellan de två kommunikationssätten troligtvis helt kommer att försvinna.

2.1.3. Digitala bibliotek

Om vetenskaplig kommunikation, och i ännu större utsträckning kunskapsorganisation, har en lång historia bakom sig, förhåller det sig annorlunda med digitala bibliotek. Det florerar många mer eller mindre synonyma ord – digitala bibliotek, elektroniska bibliotek, virtuella bibliotek, hybridbibliotek. Ofta likställs digitala och elektroniska bibliotek (Deegan & Tanner 2002, s. 20) Vad är då ett digitalt/elektroniskt, virtuellt eller hybridbibliotek? Det är lättare att definiera vad ett digitalt bibliotek ska göra - tillgängliggöra kunskap och information i digital form med hjälp av modern informationsteknik - än att definiera vad det digitala biblioteket är.

Ofta är arbetet med att göra elektroniska resurser tillgängliga integrerat i ett fysiskt biblioteks arbete med att tillgängliggöra pappersbaserade dokument och på senare år även andra medietyper såsom CD, videofilm och multimedia. Denna typ av bibliotek, där tryckt och elektroniskt material existerar sida vid sida, brukar ofta kallas hybridbibliotek (Deegan & Tanner 2002, s. 21). Owen (2000) väljer att kalla hybridbiblioteket för ”a ’hybrid’ model of information services”, och betonar därmed att det inte bara är innehållet i biblioteket som har förändrats, utan även formen för verksamheten. Han tror att ”hybridmodellen” bara kommer att finnas under en övergångsfas och ganska snart ersättas av en ”virtuell modell”, åtminstone inom vetenskaplig kommunikation.

Det virtuella biblioteket kan ses som det fulländade digitala biblioteket, biblioteket utan väggar (Deegan & Tanner 2002, s. 20), som enbart hanterar digitala dokument vilka blir tillgängliga för användaren över ett nätverk. Owen (2000) tror att tryckta dokument kommer att förstås existera, men inte inom samma system som de digitala. Tryckta dokument kommer dessutom bara att ”spela en ytterst begränsad roll, åtminstone inom ett antal vetenskapliga områden” (Owen 2000; *min översättning*). Owen menar också att det med hjälp av ny informationsteknik kommer att utvecklas ”nya modeller för kunskapsrepresentation, distribution och återvinning” inom vetenskaplig kommunikation (2000; *min översättning*).

Deegan & Tanner (2002) gör en mindre genomgång av olika definitioner av det digitala biblioteket. De bidrar även med en egen definition av vad som karakteriserar det digitala biblioteket, men en definition mer i förhållande till andra typer av digitala samlingar, t.ex. webben, än i förhållande till det fysiska biblioteket.

1. A digital library is an managed collection of digital objects.
2. The digital objects are created or collected according to principles of collection development.
3. The digital objects are made available in a cohesiv manner, supported by services necessary to allow users to retrieve and exploit the resources just as they would any other library materials.
4. The digital objects are treated as long-term stable resources and appropriate processes are applied to them to ensure their quality and survivability. (Deegan & Tanner 2002, s. 22)

I ovanstående definition läggs tyngdpunkten på innehållet och viss mån på själva återvinningssystemet i det digitala biblioteket. Detta är, enligt Borgman (2000a, s. 415), det sätt på vilket dataspecialister/informationsvetare vanligtvis definierar det digitala biblioteket. Bibliotekarier, å andra sidan, brukar i högre grad definiera det digitala biblioteket utifrån den institution eller organisation som står bakom det digitala biblioteket. Borgman har i en tudelad definition försökt spegla det digitala biblioteket utifrån ett både informations- och biblioteksvetenskapligt synsätt.

1. Digital libraries are a set of electronic resources and associated technological capabilities for creating, searching and using information. In this sense they are an extension and enhancement of information storage and retrieval systems that manipulate digital data in any medium (text, images, sounds; static or dynamic images) and exist in distributed networks. The content of digital libraries includes data, metadata that describe various aspects of the data (e.g. representation, creator, owner, reproduction rights), and metadata that consist of links or relationships to other data or metadata, whether internal or external to the digital library.
2. Digital library are constructed – collected and organized – by [and for] a community of users, and their functional capabilities support the information needs and uses of that community. They are a component of communities in which individuals and groups interact with each other, using data, information and knowledge resources and systems. In this sense they are an extension, enhancement, and integration of a variety of information institutions as physical places where resources are selected, collected, organized, preserved, and accessed in support of a user community. These information institutions include, among others, libraries, museums, archives and schools, but digital libraries also extend and serve other community settings, including classrooms, offices, laboratories, homes and public spaces. (Borgman 2000a, s.415)

Enligt min åsikt är Borgmans definition att föredra, jämfört med en definition som begränsar det digitala biblioteket till enbart innehåll och olika tekniker för informationsåtervinning.

2.2. Teoretisk referensram

Som en teoretisk referensram för min uppsats kommer jag att använda mig av några olika texter av fyra författare. Ingen av dessa texter berör direkt OAI, men kan enligt min mening användas som en utgångspunkt för en studie av OAI: s återvinningsmodell utifrån ett historiskt, socialt och kulturellt perspektiv. De fyra författarna är Jack Andersen, Christine L. Borgman samt författarparet John Seely Brown och Paul Duguid.

Andersen är anställd vid Institutet för Informationsstudier vid Danmarks Biblioteksskola i Köpenhamn. Han skriver f.n. på sin doktorsavhandling, en ”teoretisk analys av kunskapsorganisationens roll inom vetenskaplig kommunikation” (Andersen 2002, s. 29; *min översättning*). Den text av Andersen som jag kommer att referera till är en artikel publicerad i tidskriften Knowledge Organization, med titeln *Communication technologies and the concept of knowledge organization - a medium-theory perspective* (Andersen 2002).

Andersen menar att den gängse definitionen av begreppet kunskapsorganisation är allt för snäv. Kunskapsorganisation bör, hävdar Andersen, sättas in i ett socialt och historiskt sammanhang. Han menar att ”kunskapsorganisation som begrepp är underordnat och interagerar med ett samhälles sociala kunskapsorganisation” (Andersen 2002, s. 29; *min översättning*). Begreppet social kunskapsorganisation innebär att ett samhälle organiserar sin kunskap i bl.a. lärda institutioner, tidskrifter och bibliotek (Hjörland 1997, s. 45). Ett tydligt exempel på denna sociala kunskapsorganisation är samhällets indelning av vetenskaplig kunskapsproduktion i olika discipliner (Burke, P. 2000, s.81ff). Resultatet av den sociala kunskapsorganisationen är bl.a. dokument, vilka sedan organiseras utifrån olika kunskapsorganisatoriska principer (Andersen 2002, s. 32).

Enligt Andersen har ett antal forskare inom Biblioteks- och informationsvetenskap, t.ex. Joakim Hansson (1999) och Birger Hjörland (1998), under senare år diskuterat kunskapsorganisation utifrån historiska, kulturella och sociala aspekter. Få har dock, menar Andersen, tagit ett ”kommunikations- eller mediaperspektiv i beaktande” (2002, s. 29; *min översättning*). Andersen försöker råda bot på detta genom att utifrån ett medium-teoretiskt perspektiv analysera begreppet kunskapsorganisation. Hans utgångspunkt är Meyrowitz Medium Theory. Enligt denna teori påverkar olika kommunikationstekniker vad och hur människor kommunicerar och ny kommunikationsteknik kan i på sikt förändra ett samhälles sociala organisation. Andersen menar att en sådan förändring även får konsekvenser för hur ett samhälle väljer att organisera sin kunskap (Andersen 2002, s. 31).

Utifrån hypotesen att ny kommunikationsteknik i grunden förändrar vårt sätt att tänka och ”producera, kommunicera och organisera kunskap” (2002, s. 29; *min översättning*), analyserar Andersen begreppet kunskapsorganisation utifrån fyra olika kommunikationstekniker och hur dessa genom tiderna förändrat människors sätt att kommunicera. De fyra kommunikationsteknikerna är tal, skrift, tryckta och elektroniska

medier. I kulturer med enbart muntlig kommunikation är ”kunskapsorganisation och social organisation sammanvävda därför att i dessa kulturer existerar inga externa sätt att lagra kunskap” (Andersen 2002, s. 37; *min översättning*). Med kommunikationsteknikerna skrift, tryckpressar och datorer ser det annorlunda ut:

Writing, printing and computing have, throughout the history, brought about significant social and cultural changes in modern societies that have resulted in changes in the social organization of knowledge in society. They have done so by fostering new ways of producing, communicating, storing, retrieving, and organizing knowledge. Writing, printing, and the computer/internet have in this connection all contributed to changes in the LIS concept of knowledge organization. (Andersen 2002, s. 37)

Andersen menar att det är nödvändigt att betrakta kunskapsorganisation utifrån ett brett socialt och historiskt perspektiv, inte minst när det gäller att förstå ”Internet som en växande källa för kunskap och information” (Andersen 2002, s. 29; *min översättning*). Utifrån Finnemans tankar beskriver Andersen (2002, s.35f) hur datorer och i ännu större utsträckning Internet, i dess egenskap av ”globalt distribuerat, elektroniskt integrerat arkiv av kunskap” (Finneman 1999a, s. 9; *min översättning*) leder till att kunskapsorganisation mer än någonsin blir en del av "det som kan kallas 'ett samhälles sociala, epistemologiska infrastruktur'" (Andersen 2002, s. 36; *min översättning*). Här kan man, anser jag, se en koppling mellan Andersen och OAI: s arbete, vars slutmål är just ett ”kunskapsarkiv” med världsomfattande täckning.

På vilket vad sätt har då Andersens tankar någon relevans för min uppsats? Utvecklingen av ”öppna arkiv” och andra publiceringsalternativ, vilka kan bli en realitet just genom nya kommunikationstekniker, kommer utan tvekan att på ett djupgående sätt förändra forskares sätt att kommunicera med varandra och med omvärlden. Denna förändring i forskares sätt att kommunicera, förändrar även det sätt på vilket vetenskaplig information lagras, söks och återvinns. Något som Andersen inte berör i sin artikel, men som andra debattörer påpekar, är att datorer och Internet inte har förändrat alla discipliners sätt att kommunicera i samma utsträckning. En av dem är Christine L. Borgman.

Borgman arbetar vid University of California, avdelningen för Information studies. White & McCain (1998) har med hjälp av en co-citeringsanalys undersökt hur disciplinen Library and Information Science (LIS) utvecklats under perioden 1972-1995. Deras resultat visar hur olika delområden inom LIS utvecklats i förhållande till varandra och vilka forskare som dominerar inom vilka delområden. Enligt undersökningen är Borgman mest verksam inom följande tre delområden: återvinning on-line, användarteori och OPACs (White & McCain 1998, s. 335). Undersökningen visar också att Borgman är en forskare med allt större inflytande inom LIS. Under senare år har hon skrivit både monografier och artiklar om ny informationsteknik, digitala bibliotek, vetenskaplig kommunikation, m.m. Av titlarna att döma skriver hon ofta om dessa frågor utifrån ett socialt perspektiv, där mötet mellan människa och teknik spelar en viktig roll.

Den artikel av Borgman som jag tänker använda som teoretisk utgångspunkt heter *Digital libraries and the continuum of scholarly communication*. I artikeln undersöker Borgman relationen mellan digitala bibliotek och vetenskaplig kommunikation, eller

snarare på vilka sätt forskning inom dessa två områden tangerar varandra. Borgman (2000a, s. 414) beskriver hur mötet mellan dessa två forskningsområden kan studeras utifrån både processer och strukturer inom vetenskaplig kommunikation. Jag kommer att begränsa mig till den del av Borgmans artikel som behandlar relationen digitala bibliotek och processer inom vetenskaplig kommunikation.

Enligt Borgman finns det tre intressanta forskningsområden relaterade till ”skärningspunkten mellan processer inom vetenskaplig kommunikation och digitala bibliotek” (2000a, s. 412; *min översättning*). Det första av dessa tre områden är det digitala bibliotekets förmåga att bistå forskare i informationscykeln. I denna cykel ingår sökning och användning samt skapande av information/kunskap. Dessa tre aktiviteter utfördes tidigare separat, men med hjälp av ny teknik sker arbetet med att söka, använda och skapa information via ett och samma medium (Borgman 2000a, s. 417). Detta har även Finneman (1999) uppmärksammat. Finneman menar att uppgifter som tidigare genomfördes med hjälp av olika medier, idag kan utföras med hjälp ett och samma medium, datorn. Finneman urskiljer tre olika funktioner. Den första inbegriper bl.a. produktion, lagring och återvinning av kunskap. Den andra inbegriper presentation av kunskap och den tredje kommunikationen av kunskap. Enligt både Borgman och Finneman påverkar den nya tekniken alla delar av ett samhälles kunskapsproduktion. Jag håller för troligt att i synnerhet forskares sätt att använda och producera kunskap kommer att påverkas, eftersom forskare är en grupp i samhället som ofta tidigt får tillgång till ny teknik. Enligt Borgman har det dock visat sig att olika forskarkollektiv använder och producerar kunskap på olika sätt. Som exempel nämner hon att forskare inom olika discipliner har skilda sätt att återanvända andras kunskap för att skapa ny kunskap. Kunskap om dessa olika förhållningssätt kan enligt Borgman användas i utvecklingen av digitala bibliotek.

Det andra forskningsområde som Borgman berör är dokumentens "sociala liv" (2000a, s. 417f; *min översättning*). Hon bygger här på Brown och Duguids (1996) tankar i artikeln *The social life of documents*. Dokument ses av vissa forskare enbart som informationsbärande ”saker”. Borgman hänvisar här till bl.a. Michael Buckland (1991). Denna definition är, enligt Borgman, allt för snäv för de dokument som används inom vetenskaplig kommunikation. Dessa dokument har, i termer av Brown och Duguids språkbruk, även ett socialt liv. Enligt Brown och Duguid har alla typer av dokument ett socialt liv, inte bara vetenskapliga dokument. Med socialt liv menar de att dokument inte bara ska ses som informationsbehållare, utan att de också skapar och vidmakthåller sociala grupperingar. Utifrån ett forskarperspektiv finns det olika skäl till hur och varför vetenskapliga dokument skapas och används, vilket innebär att ”en del av ett dokumentets mening kan finnas i ett kollektiv och en del i dokumentet självt” (Borgman 2000a, s. 417; *min översättning*). Enligt detta synsätt finns olika sociala processer införlivade i vetenskapliga dokument och Borgman menar att kunskap om dessa processer kan vara värdefulla i uppbyggnaden av digitala bibliotek.

Det tredje forskningsområde som Borgman tar upp är elektronisk publicering (2000a, s. 419f). Idag är en stor del av publikationsprocessen elektronisk och detta oavsett om de färdiga dokumenten förmedlas i tryckt eller digital form. Borgman menar att, om det enbart vore tekniken som styr, så skulle mer eller mindre alla dokument förmedlas i digital form. Fortfarande publiceras en stor del av allt vetenskapligt material i tryckt form och Borgman anser, och många med henne, att detta beror på att även andra faktorer styr. Hon hänvisar till Kling och McKim (1999, s. 897f) som i sin artikel

urskiljer tre faktorer som är avgörande för huruvida forskare anser en publiceringsform vara effektiv eller inte. Dessa tre faktorer är den valda publiceringsformens förmåga att garantera publicitet, tillförlitlighet och tillgänglighet. Än så länge tenderar dessa faktorer att "favorisera etablerade tidskrifter och förläggare som har byggt upp starka relationer och gott anseende" (Borgman 2000a, s. 420; *min översättning*). Vid uppbyggnaden av digitala bibliotek är kunskap om utifrån vilka kriterier forskare väljer publiceringsform viktig, anser Borgman.

I boken *From Gutenberg to the global information infrastructure* diskuterar Borgman (2000b) på ett mer uttömmande sätt hur viktigt hon anser det är att förstå den sociala kontext som dokument befinner sig i. Hon sammanfattar där sina tankar på följande vis:

The more we understand about the social context of documents, the better we can design electronic publications and digital libraries. We need to know more about how, why, when and for whom people create documents if we are to build better tools for them to do so. Similarly, we need to know more about peoples criteria for selecting and evaluating documents if we are to build better tools for seeking information. We need to know what people do with documents once they have them, so that we can build tools for reading, browsing, annotating, manipulating, and otherwise using documents. Hence we need to study and test implicit assumptions about information-related behaviour if we are to design better system to support that behaviour. (Borgman 2000b, s. 94)

De två sista författarna som ingår i min teoretiska referensram är John Seely Brown och Paul Duguid. De har nyligen vidareutvecklat tankegångarna från artikeln *The social life of documents* (1996), i boken *The social life of information* (2000). Brown är anställd som forskare vid Xerox, ett amerikanskt företag som bl.a. tillverkar kopieringsmaskiner. Duguid forskar inom Social and Cultural studies vid University of California. I *The social life of information* diskuterar de betydelsen av att ta hänsyn till den sociala kontext information av alla de slag befinner sig i. I boken hävdar Brown och Duguid att orsaken till att nya tekniker inte alltid får den genomslagskraft dess upphovsmän förväntar sig, kan beror på ett slags tunnelseende hos de samma. Genom att enbart fokusera på vad människan ska kunna göra med hjälp av teknik och inte på vem som ska använda den eller varför, negligeras viktiga aspekter såsom "bakgrund, historia, allmänskunskap, sociala resurser" (Brown & Duguid 2000, s. 1ff; *min översättning*).

Jag anser att Borgman och Brown och Duguid på ett bra sätt kompletterar Andersens teorier. Andersen fokuserar på hur införandet av ny kommunikationsteknik på ett djupgående sätt förändrar människans sätt att producera, kommunicera och organisera kunskap. Med tanke på hur dessa förändringar tidigare har skett genom historien, kan vi med all sannolikhet förvänta oss att IT också kommer att leda till omfattande förändringar när det gäller framställning och kommunikation av kunskap, samt i hur vi kommer att spara den för eftervärlden. Borgman och Brown och Duguid utgår också från att ny teknik påverkar människans sätt att förhålla sig till kunskap, men fokuserar i högre grad på samspelet mellan människa och teknik. Enligt mitt sätt att se är det så här vi bör förhålla oss till det som sker. Tekniken påverkar oss, men människan påverkar även tekniken.

2.3. Kunskapsöversikt

I och med att jag har valt att undersöka OAI utifrån flera olika perspektiv finns det artiklar, studier och forskningsprojekt inom ett flertal områden som skulle kunna bidra till en ökad förståelse för ”öppna arkiv” och egenarkivering. Utifrån OAI, kunskapsorganisation, vetenskaplig kommunikation och digitala bibliotek kommer jag nedan att kortfattat diskutera och redogöra för vissa artiklar, tidigare studier och forskningsprojekt med relevans för min uppsats.

En del av de nämnda artiklarna/studierna/forskningsprojekten kommer jag att återvända till och använda som empiri i kommande avsnitt. Jag kommer inte att nämna alla texter som jag har för avsikt att använda mig av längre fram, utan har valt ut några, i mitt tycke, representativa texter.

2.3.1. OAI-studier

När det gäller OAI gör jag bedömningen att det än så länge finns ganska begränsat med forskning. Resultaten av en del studier och översikter har dock publicerats under de senaste två, tre åren. Dobratz och Matthaei (2003) presenterade i januari 2003 en mindre undersökning för Open Archives Forums räkning, en intresseorganisation inriktad på att sprida kunskap om och främja olika OAI-relaterade aktiviteter inom Europa. Dobratz och Matthaei har kartlagt utvecklingen av OAI-aktiviteter inom Europa under senare år. Vissa jämförelser görs också med Amerika, Australien och Asien. Trots ett ganska begränsat underlag, totalt ingår ett drygt hundratal OAI-aktiviteter i materialet, visar resultatet något överraskande att Europa har kommit längst. Amerika kommer på en god andraplats, medan Australien och Asien än så länge har få OAI-aktiviteter. Inom Europa skiljer det stort mellan olika länder. Storbritannien har kommit längst, därefter Tyskland och på gemensam tredje plats Frankrike och Sverige.

Två artiklar, *Setting up an institutional e-print archive*, (Pinfield, Gardner & MacColl 2002) och *Open archives and UK Institutions* (Pinfield 2003) beskriver några brittiska universitets erfarenheter av att upprätta ”öppna arkiv”. Innehållet i dessa två artiklar kan inte direkt klassas som forskning, men bidrar enligt min mening med kunskap som kan bidra till en större förståelse för ”öppna arkiv”.

Jag har endast funnit två studier som undersöker forskares attityder till ”öppna arkiv”, en avslutad och en pågående. Den avslutade studien (Hunt 2002) jämför attityder mellan användare, respektive ickeanvändare, av två olika ”öppna arkiv”, det ena inom fysik och det andra inom kognitiva vetenskaper. Den pågående studien (Gadd, Oppenheim & Proberts 2003) undersöker hur forskares attityder gentemot ”öppna arkiv” påverkas av olika upphovsrättsliga frågor.

2.3.2. Kunskapsorganisation

Det finns, utan tvekan, många forskningsområden inom kunskapsorganisation som har relevans för utvecklingen av ”öppna arkiv”. Ett viktigt område är forskning kring metadata, eftersom OAI:s modell bygger på utbyte av metadata enligt ett standardiserat format. I avsnitt 4.5 kommer jag att gå igenom vad metadata är och hur OAI-kompatibla arkiv fungerar. Här redovisar jag därför bara helt kort varför jag anser att forskning kring metadata är viktig.

Metadata kan kort beskrivas som mer eller mindre utförliga beskrivningar av digitala dokument, d.v.s. en slags katalogposter vars syfte är att identifiera och underlätta organisation, sökning, urval, åtkomst och utbyte av dokument. Dessa beskrivningar skapas ofta av de digitala dokumentens författare, till skillnad från katalogposter för t.ex. monografier som skapas av utbildade katalogisatörer. Utbyte av metadataposter är en av grundstenarna i OAI: s återvinningsmodell.

Ett exempel på forskning med anknytning till skapande av metadata är Metadata Generation Research Project vid School of Information and Library Science, i North Carolina, USA. Genom projektet vill forskarna få kunskap om effektiva sätt att skapa metadata. De har i olika undersökningar studerat dels författargenerad metadata (Greenberg et al. 2001), dels metadata skapad utifrån ett samarbete mellan författare och erfarna katalogisatörer (Greenberg & Davenport Robertsson 2002). Dessa två undersökningar är inte helt överensstämmande när det gäller resultat. Greenberg et al. (2001) visar att författare utan problem kan skapa metadata av god kvalitet, medan Greenberg och Davenport Robertsson (2002) snarare talar för ett samarbete mellan författare och experter, för att höja kvaliteten på metadata. Jag håller det för troligt att det inom olika discipliner finns varierande kunskaper om metadata - varför metadata behövs och hur den bör utformas. Bibliotekarier vid de digitala bibliotek som har eller ämnar skapa "öppna arkiv" skulle troligtvis vara behjälpta av mer kunskap om dessa olikheter.

Tidigare undersökningar av hur stor andel av dokument publicerade på webben som innehåller metadata visar på ganska nedslående siffror. En undersökning, av Lawrence & Giles från 1999 (se Smith 2002, s.135), visade att endast 0.3 % av alla dokument på webben hade metadata enligt det format som används inom OAI, det s.k. Dublin Core-formatet. Tre år senare visade Smith (2002, s.35) att andelen webbsidor med metadata enligt Dublin Core hade ökat till 4.35 %, inte direkt någon lysande utveckling. Dessa siffror tyder på att det finns problem förknippade med skapandet av metadata och Smith efterfrågar mer forskning om metadata i sig och effekterna av kvalitativa och kvantitativa förändringar av metadata för webbresurser. Denna typ av forskning skulle troligtvis kunna vara till nytta även för aktiviteter inom OAI: s verksamhet.

2.3.3. Vetenskaplig kommunikation

Under efterkrigstiden expanderade forskning, vetenskaplig publicering och biblioteks- och universitetsverksamhet kraftigt. Detta ledde till ett ökat intresse för fenomenet vetenskaplig kommunikation. En av pionjärerna var Jack Meadows (1974), som i sin bok *Communication in science* bl.a. diskuterar hur och varför forskare inom naturvetenskap forskar, kommunicerar och publicerar sig. Under 1980-talet och en bit in på 1990-talet var forskningen kring vetenskaplig kommunikation mindre intensiv, för att sedan i mitten av 1990-talen åter bli aktuell (Borgman 2000a, s. 412f).

Att vetenskaplig kommunikation återigen har kommit i fokus beror delvis på IT-utvecklingen under de senaste 10-15 åren. Datorer, Internet, webben, tillkomsten av digitala bibliotek och övergången från tryckt till elektronisk publicering, har i stor utsträckning förändrat förutsättningarna för forskare, både vad gäller forskningen i sig och hur den kommuniceras (Borgman 2000a, s.413). Som en delvis följd av dessa nya kommunikationssätt genomgår även den traditionella publiceringskedjan en förändring

av sällan skådat slag, vilket ger ytterligare incitament till att forska kring vetenskaplig kommunikation.

Som tidigare nämnts menar Borgman att det finns det två större forskningsområden inom vetenskaplig kommunikation (2002a, s.414). Forskare inom det ena området tenderar att fokusera på sociala processer, t.ex. hur och varför forskare publicerar sig. Hit hör t.ex. Meadows och Borgman själv. Andra forskare väljer att lägga tyngdpunkten på sociala strukturer. Det kan vara forskare som genom bibliometriska studier försöker visa på relationer mellan forskare inom ett ämnesområde, t.ex. White & McCain (1998) eller forskare som studerar vetenskaplig kommunikation som en del av en struktur, där även universitet, bibliotek och förlag ingår (Borgman 2000a, s.414). Det är inte lätt, menar Borgman (2000a, s.424), att definiera var gränsen går mellan sociala processer och sociala strukturer inom vetenskaplig kommunikation, även om forskare som sagt tenderar att fokusera den ena eller den andra aspekten. Ett författarpär som jag kommer att återvända till är Rob Kling och Geoffrey McKim (1999; 2000), som enligt min åsikt diskuterar både utifrån ett process- och ett strukturperspektiv.

Två artiklar som på ett tidigt stadium bidrog till att öka mina kunskaper om krisen inom vetenskaplig publicering var Hagerlid (2002) och Rabow (2001). Genom dessa två texter kunde jag förhållandevis snabbt och enkelt sätta mig in i ett problem som för mig var alldeles nytt.

Tidigare nämnda Meadows studerar som sagt processer inom vetenskaplig kommunikation. I boken *Communicating research* (1998) undersöker han hur forskare inom naturvetenskap, samhällsvetenskap och humaniora kommunicerar med varandra. Jag kommer att använda mig av Meadows monografi för att sätta in vetenskaplig kommunikation i en social och historisk kontext.

En annan forskare vars resultat har relevans för min uppsats är Ylva Lindholm-Romantschuk (1996; 1998). Enligt Lindholm-Romantschuk (1998 s.vii) riktas en större del av forskningen inom vetenskaplig kommunikation mot kommunikation inom naturvetenskapliga discipliner. Hon har därför i sina studier valt att studera informationsflöden inom samhällsvetenskap och humaniora.

Sely Costa och Jack Meadows (2000) har i en undersökning studerat en grupp ekonomer och en grupp sociologers inställning till och användning av ny teknik (IT). Undersökningen syftade till att skapa förståelse för hur ny teknik påverkar olika kommunikationsmönster bland samhällsvetenskapligt inriktade forskare.

The Knight Higher Education Collaborative (2002) redogör för en konferens om vetenskaplig kommunikation inom humaniora och samhällsvetenskap. Deras artikel bidrar bl.a. med en bild av samhällets syn på humaniora och samhällsvetenskap. Denna artikel har jag valt att ta med därför att det är en av de få artiklar jag har funnit, som diskuterar elektronisk publicering utifrån ett samhällsvetenskapligt och humanistiskt perspektiv.

2.3.4. Digitala bibliotek

Digitala bibliotek är ett förhållandevis nytt fenomen. Detta innebär att det är först under de senaste 10 åren som man har börjat forska kring digitala bibliotek, även om det har

forskats inom relaterade ämne under betydligt längre tid än så (Borgman 2000b, s.34f). Svårigheten att definiera det digitala biblioteket återspeglas i viss mån inom forskningen, menar Borgman (2000a s. 424). Sedan några år tillbaka har det i USA beviljats stora anslag till forskning kring digitala bibliotek, vilket enligt Borgman lockat forskare från andra områden än biblioteksområdet, alla med olika fokus. Dessa olika forskares intresseområden återspeglar även ämnets tvärvetenskapliga natur (Borgman 2000b, s.36). Forskare med inriktning på datavetenskap diskuterar främst tekniska frågor, forskare från biblioteks- och informationsvetenskap (LIS) t.ex. innehåll och organisation, sociologer den sociala kontexten och ekonomer ekonomiska modeller, etc. Samarbetet disciplinerna emellan fungerar, enligt Borgman, inte helt friktionsfritt och hon efterlyser ett ökat samarbete och större ömsesidig respekt, i synnerhet mellan forskare inom datavetenskap och LIS.

Med dessa rader avslutar jag den teoretiska delen av min uppsats. I följande avsnitt redogör jag för val av metod och hur jag har gått tillväga för att finna information.

3. Metod

3.1. Metodval

Jag har som metod valt att göra en litteraturstudie. Urval av litteratur kommer att se något olika ut för mina två syften. När det gäller det första syftet, att undersöka och beskriva OAI utifrån ett rent kunskapsorganisatoriskt perspektiv, kommer jag att använda mig av dokument *om* OAI och dokument publicerade *av* organisationen OAI och andra aktörer inom alternativ publicering. Det finns en hel del material om ämnet, varav övervägande delen är elektroniskt publicerade artiklar. Många av artiklarna har varit överraskande lika när det gäller innehåll. Troligtvis beror detta på att företeelsen är förhållandevis ny och många av de artiklar jag har läst är beskrivningar, snarare än analyser av OAI, avsedda att informera olika läsekretsar om en ny företeelse inom vetenskaplig kommunikation. I de artiklar som har publicerats under 2002-2003 tycker jag mig kunna se en något mer analyserande och kritisk attityd, även bland dem som i grunden är positiva till OAI och andra alternativ. Det tyder på att fenomenet egenarkivering ”börjar få fäste”, åtminstone inom vissa delar av den akademiska världen.

Efterhand som jag läste in mig på ämnet har jag kunna urskilja ett antal författare som förekommer mycket flitigt i debatten och det är tydligt att det finns en kärna av debattörer som arbetar *för* en snabb omstrukturering av den vetenskapliga kommunikationen. Kritikerna är som sagt färre. Uppsatsens avsnitt om OAI kommer till stor del att bli en beskrivning liknande den som finns i många av de artiklar jag har läst. Enligt min bedömning finns det inte sedan tidigare någon sådan beskrivning på svenska och jag anser därför att den har en funktion att fylla. Min avsikt är dock att, utifrån de artiklar jag har läst, syna OAI med så kritiska ögon som möjligt.

För att uppnå mitt andra syfte, att närma mig OAI och vetenskaplig kommunikation utifrån ett socialt, kulturellt och historiskt perspektiv, kommer jag att använda mig av delvis annorlunda litteratur. Här förekommer även några monografier. Jag kommer att använda mig av litteratur som ibland, men inte alltid, direkt berör OAI. En del av litteraturen handlar mer om vetenskaplig kommunikation i allmänhet och då i synnerhet

om vetenskaplig kommunikation via tryckta medier. Utbudet av litteratur om vetenskaplig kommunikation är, jämfört med litteratur med direkt anknytning till OAI, enormt. Jag har därför fått begränsa mig till ett fåtal författare med vars hjälp jag avser att sätta in OAI i en social, kulturell och historisk kontext. En del av texterna är av mer teoretisk karaktär, andra rena redovisningar av olika empiriska undersökningar. Dessa undersökningar blir, tillsammans med de teoretiska texterna, min empiri, d.v.s. det material med vars hjälp jag ska besvara mina frågeställningar.

I inledningen av mitt uppsatsarbete hade jag funderingar på att göra en egen empirisk undersökning, t.ex. djupintervjuer av ett mindre antal forskare inom Biblioteks- och informationsvetenskap eller en mer allmänt hållen attitydundersökning riktad mot ett större antal forskare. Efterhand kom jag till insikt om att en sådan undersökning, för att kunna generera tillräckligt omfattande och intressanta svar, skulle behöva vara långt mer omfattande än vad tiden medgav. Dessutom tror jag att det hade varit svårt att på ett konstruktivt sätt undersöka svenska forskare attityder till något som är så förhållandevis nytt som ”öppna arkiv”. Men, ett flertal universitet runt om i landet upprättar nu på försök olika alternativ för vetenskaplig kommunikation och tiden kan snart vara mogen för en attitydundersökning.

3.2. Informationssökning

Mitt uppsatsämne inbegriper flera olika problemområden - ”öppna arkiv”, kunskapsorganisation, vetenskaplig kommunikation och digitala bibliotek. Jag tog min utgångspunkt i texter som handlade om ”öppna arkiv”. Med tanke på att ”öppna arkiv” anknyter till elektronisk publicering, föll sig valet av källa ganska naturligt – webben. Jag hittade information i elektroniska tidskrifter och på olika organisationers och universitets hemsidor. Det finns många organisationer och universitet som är involverade i utvecklingen av ”öppna arkiv”. De jag refererar till är några av dem som oftast omnämns i den litteratur jag har läst. Jag förstod ganska omgående att det här är ett område där utvecklingen går oerhört snabbt och för att få tillgång till den mest uppdaterade informationen framstod webben som det bästa alternativet. Ett problem med ett flertal av de texter jag hittade på webben var att de förhöll sig tämligen okritiska. Många av texterna var också, som tidigare nämnts, snarlika till innehåll, även om jag har kunnat se en tendens till en, om inte ifrågasättande, så mer kritisk hållning på senare tid. Via referenslistor i olika artiklar och papers på webben letade jag mig vidare, bakåt i tiden och ut från webben. Det visade sig finnas en del skrivet även i tryckta vetenskapliga tidskrifter, men utbudet på webben tycks än så länge långt mer omfattande. Med tanke på ämnets karaktär, håller jag för troligt att dominansen av digitala artiklar kommer att bestå.

Jag har även sökt i databaserna *Library and Information Sciences Abstracts*, *Social Services Abstracts*, *Sociological Abstracts* och *Library Literature and Information Science*. Via dem fann jag både elektroniska och tryckta källor med relevans för min uppsats. Biblioteket vid Högskolan i Borås har naturligtvis också bidragit med en del material, såväl tidskrifter som monografier. Databaser använde jag mig huvudsakligen av i början av mitt uppsatsarbete. Under den senare delen har jag mestadels hittat ny litteratur genom andra artiklars och monografiens referenslistor. Risken med ett sådant förfarande är att man blir kvarhållen inom en krets av ”likasinnade”. Detta blev mycket tydligt efter att jag hade läst ett otal texter om ”öppna arkiv”, där samma författare och samma texter dök upp gång på gång, i en slags rundgång. Via en av de få något mer

kritiska texterna lyckades jag ta mig över till ”den andra sidan”, där de mer ifrågasättande texterna fanns och hoppas på så sätt kunna ge en någorlunda balanserad bild av problemområdet.

Det har inte varit några problem att hitta litteratur, tvärtom. Efter ett tag uppstod en litteraturövermättning och jag insåg då att jag måste börja begränsa mig, för att överhuvudtaget kunna hantera materialet. För att hålla mig à jour med utvecklingen har jag dock med jämna mellanrum tittat igenom två av de elektroniska tidskrifter som skriver mycket om alternativ publicering, *Ariadne* och *D-Lib Magazine*.

4. Open Archives Initiative

I detta avsnitt kommer jag att undersöka hur ”öppna arkiv” fungerar utifrån ett grundläggande kunskapsorganisatoriskt perspektiv, vilket är uppsatsen första syfte. Jag kommer dock att börja med en beskrivning av den vetenskapliga tidskriftens bakgrund och även ge en mer nutida bakgrund till framväxten av OAI. Denna historiska inledning gör att följande avsnitt även delvis hör ihop med uppsatsens andra syfte, att undersöka ”öppna arkiv” utifrån ett historiskt, socialt och kulturellt perspektiv.

4.1. Historisk bakgrund

Den grundläggande tanken med OAI: s arbete är att underlätta och finna nya former för vetenskaplig kommunikation, med hjälp av modern informationsteknik. Vetenskaplig kommunikation i sig är dock inte ett nytt fenomen. Så länge det har funnits vetenskapsmän har dessa personer haft behov att kommunicera med varandra och omvärlden. Till en början skedde detta utbyte av idéer genom "personliga kontakter, brev och informella nätverk" (Rabow 2001, s. 34). Under Renässansen formaliserades kommunikationen i det som kom att kallas lärda sällskap (Rabow 2001, s. 34), t.ex. *the Royal Society* i London. De brev som personerna i dessa lärda sällskap skickade till varandra har av Brown och Duguid (1996) beskrivits som "prototypen för vetenskapliga tidskrifter".

De lärda sällskapen uppstod under en tid av stora förändringar. I den vetenskapliga revolutionens spår ersattes skolastiken av en "praktisk och nyttoinriktad filosofi och vetenskap" (Bacon 1990, s. 202f), med en ny syn på hur kunskap skulle erövrats. En av dem med störst inflytande under denna period var Francis Bacon, 1561-1626. Bacon "betonade särskilt den betydelse den nya naturvetenskapen skulle få för mänskligheten, ..." (Bacon 1990, s. 202f) och förespråkade systematiska, empiriska vetenskapliga undersökningar där vetenskapsmän skulle bygga vidare på varandras forskningsresultat. För att detta skulle bli möjligt måste alla forskningsresultat organiseras och spridas, t.ex. genom akademier och tidskrifter, för att eventuella misstag inte skulle göras om (Savanije & Smith 2002, s. 1). En ny kunskapssyn, i kombination med att antalet vetenskapsmän steg kraftigt, sägs ha lett fram till uppkomsten av ett nytt medium för vetenskaplig kommunikation, den vetenskapliga tidskriften. I grunden låg naturligtvis också ny teknik. I Europa var boktryckarkonsten vid den här tiden drygt 100 år gammal. Vetenskapliga rön hade också redan tidigare publicerats i bokform. Som publiceringsform ansågs böcker ha vissa nackdelar jämfört med tidskrifter, framförallt för att det tog så lång tid att producera böcker. Men också för att den nya tidens vetenskapliga tankeutbyte krävde ett mer dynamiskt medium än boken, som till sin karaktär var mer definitiv. Vetenskapsmännen sökte en alternativ plattform till brev och

böcker och ur detta behov skapades den vetenskapliga tidskriften (Savanije & Smith 2002, s.1f). Två olika tidskrifter sägs vara de första föregångarna till vår tids vetenskapliga tidskrifter, Philosophical Transactions utgiven av the Royal Society och den franska tidskriften Journal des Scavans. Båda publicerades första gången 1665, med några månaders mellanrum (Guédon 2001).

Genom de vetenskapliga tidskrifterna kunde vetenskapsmän på ett mer formellt sätt sprida ny kunskap. Men vetenskapliga tidskrifter, i egenskap av dokument, behöver inte enbart ses som en slags bärare av idéer eller informationsbehållare (Brown & Duguid 1996). Dokument har även, hävdar Brown och Duguid (1996), ett socialt liv och med det menar de att dokument skapar och vidmakthåller sociala grupperingar. Den vetenskapliga tidskriften är den dokumenttyp som "binder samman intellektuella grupperingar" (Brown & Duguid 1996; *min översättning*). I skenet av dessa tankar ska jag redogöra för de olika funktioner den tryckta, vetenskapliga tidskriften anses fylla, utöver att sprida kunskap. Rowland (1997) nämner kvalitetskontroll, arkivering av vetenskapliga resultat och erkännande av författare, även kallat registrering av vetenskaplig prioritet. Till dessa fyra funktioner lägger Hagerlid (2002, s. 99) funktionen statusmarkör, som får anses hänga tätt samman med etablering av vetenskaplig prioritet. Vilken av dessa fyra, fem funktioner som är viktigast, finns det troligtvis skilda åsikter om. Både Rowland (1997) och Solomon (2002) hävdar att den vetenskapliga tidskriftens viktigaste funktion aldrig har varit att sprida kunskap, utan att det snarare handlar om att etablera vetenskaplig prioritet. Alla vetenskapsmän är kanske inte benägna att hålla med Solomon och Rowland om att så är fallet, men otvivelaktigt är det så att publicering i vetenskapliga tidskrifter har stor betydelse för en forskares status i den akademiska världen, både när det gäller forskningsanslag och tillsättning av tjänster, vilket bl.a. Rabow (2001, s. 35) nämner.

Funktionen kvalitetskontroll figurerar ofta i debatten kring "öppna arkiv" och kräver sin förklaring. Efter andra världskriget utvecklades inom förlagen ett referentsystem, bestående av en redaktion och ett antal referenter, s.k. *peer-reviewers*. Dessa referenter (sakkunniga forskare) ska opartiskt bedöma och välja ut vilka artiklar som ska publiceras. Peer review är tänkt att fungera som ett slags kvalitetsfilter för forskningen (Rodman 2003). Även om tanken med en strikt, opartisk och rättvis kvalitetsbedömning kan tyckas god, är peer review i sig inte förskonat från kritik. Nackdelar med det nuvarande systemet är, enligt Rodman (2003), att processen inte dokumenteras, sker i bakvänd ordning; review kommer före publicering och att det finns risk för partiskhet och korruption. Det råder inget tvivel om att kvalitetskontroll alltid kommer att vara viktig inom forskning. Däremot pågår det en intensiv debatt om hur denna kvalitetskontroll bör gå till, när en allt större del av forskningsresultaten arkiveras/publiceras via andra kanaler än den traditionella vetenskapliga tidskriften.

Under mer än 300 år behöll den vetenskapliga tidskriften sin särställning som förmedlare av vetenskaplig kommunikation, i synnerhet inom naturvetenskap, medicin och teknik. Under andra hälften av 1900-talet skedde en rad förändringar inom forskning och vetenskaplig kommunikation, förändringar som alla kan sägas leda fram till det som ibland kallas ett paradigm- eller systemskifte inom vetenskaplig kommunikation. Efter andra världskrigets slut började forskning inom främst medicin, naturvetenskap och teknik att expandera kraftigt och nya discipliner växte fram. Allt fler forskare ville publicera sig och de lärda sällskapen fick problem med att svara upp till forskarnas behov. Flera kommersiella förlag etablerade sig som publiceringsalternativ

till de lärda sällskapen. Nya tidskrifter skapades, vilket ledde till att "forskningen 'balkaniserades' genom att rapporteringen splittrades på allt fler tidskrifter" (Rabow 2001, s. 35). Redan i början av 1970-talet inleddes den prisspiral som har fortsatt ända fram till idag. De kommersiella förlagen drevs naturligtvis av ett vinstintresse. Men man kan också se att det hade utvecklats en "mycket prisokänslig efterfrågan" (Hagerlid 2002, s. 97). Genom citeringsindexet Science Citation Index, skapat av Eugene Garfield i slutet av 1960-talet, fick vissa tidskrifter högre status än andra. De mest citerade tidskrifterna kom att utgöra de s.k. kärntidskrifterna. För forskarna blev det viktigt att publicera sig i tidskrifter med hög status och för biblioteken blev det viktigt att ha kärntidskrifterna i sin samling, till i stort sett vilket pris som helst (Hagerlid 2002, s. 97f).

Under senare år har allt fler börjat kritisera den rådande situationen. För universitetsbibliotekens del har de ständiga prisökningarna inneburit att kostnaderna för vetenskapliga tidskrifter tar en allt större del av bibliotekens budget. Detta, i sin tur, leder till att biblioteket får dra in på andra områden, som exempel nämner Rabow (2001, s. 35) att inköpen av monografier har minskat kraftigt. Det är dock inte enbart rent ekonomiska aspekter som har fått biblioteks- och forskarvärlden att höja sina röster. En minst lika bidragande orsak är att, så som publiceringssystemet fungerar idag begränsas spridningen av forskningsresultat (Hagerlid 2002, s. 98). För att råda bot på denna situation har under senare år en rad alternativ för arkivering och publicering initierats, varav OAI är ett.

OAI är den aktör som arbetar för att finna tekniska lösningar för egenarkivering och samsökning av "öppna arkiv". Det finns ett flertal andra aktörer som på andra sätt stöder alternativa publiceringsformer. Jag kommer nedan att redogöra för några av de aktörer som oftare än andra omnämns i den litteratur jag har läst, innan jag går över till OAI och den återvinningsmodell som initiativet har arbetat fram.

4.2. Alternativ publicering – olika aktörer

4.2.1. Scholarly Publishing and Academic Resources Coalition (SPARC)

SPARC grundades 1998 och har sitt ursprung inom ARL, Association of Research Libraries, i USA. SPARC har över 200 medlemmar i stora delar av världen, bland dessa finns bl.a. forskningsinstitutioner och bibliotek. Genom att hjälpa forskare och lärda sällskap att starta lågpristidskrifter, hoppas man på att främja konkurrensen på marknaden för vetenskaplig kommunikation. SPARC ser också opinionsbildning som en av sina uppgifter. (Scholarly Publishing and Academic Resources Coalition 2003; Rabow 2001, s. 41; Hagerlid 2002, s.100)

4.2.2. Public Library of Science (PLoS)

PLoS grundades i oktober 2000 och är en icke vinstdrivande sammanslutning av forskare som verkar för att fritt tillgängliggöra vetenskapliga artiklar inom naturvetenskap och medicin, som tidigare har publicerats i traditionella tidskrifter. Artiklarna ska, efter den ursprungliga publiceringen, tillåtas arkiveras och göras fritt tillgängliga via nätet, i ett slags "forskningsbibliotek on-line". Ett exempel på ett sådant bibliotek on-line är PubMed Central, som arkiverar artiklar inom biomedicin. PLoS har

i december 2002 mottagit ett bidrag på 9 miljoner dollar och för dessa pengar avser PLoS att starta ett eget förlag någon gång under 2003 (Public Library of Science 2003).

4.2.3. *Budapest Open Access Initiative (BOAI)*

BOAI bildades vid ett möte i Budapest i december 2002 och är ett försök att få till stånd ett samarbete mellan alla de olika aktörer som verkar för att göra forskningsresultat mer allmänt tillgängliga. BOAI vill nå målet "open access" genom att undanröja olika barriärer, i synnerhet ekonomiska barriärer. Den litteratur som BOAI vill fritt tillgängliggöra, är i första hand granskade tidskriftsartiklar men även icke granskade artiklar, s.k. pre-prints (BOAI 2003). BOAI rekommenderar två olika strategier för att nå målet "open access". Den ena strategin är forskares egenarkivering av granskade tidskriftsartiklar i OAI-kompatibla arkiv, den andra är skapandet av en ny typ av tidskrifter, "open-access journals" (Budapest Open Access Initiative 2002).

4.2.4. *FIGARO*

FIGARO bildades i oktober 2002 och initiativtagare var två holländska och två tyska universitet. Målet är att skapa en europeisk "infrastruktur för akademisk e-publicering" (Savanije 2002, s.1; *min översättning*) och att stödja uppbyggnaden av "öppna arkiv". FIGARO vänder sig både till forskare som vill skapa ämnesarkiv och till universitet som vill skapa institutionsarkiv, men även till mindre förlag som är intresserade av alternativa publiceringsmodeller (FIGARO 2003). Den infrastruktur FIGARO bygger upp ska tillåta "en gradvis övergång från traditionell publicering till högt sofistikerade modeller" (Savanije 2002, s.5; *min översättning*).

4.3. **E-printarkiv – om bakgrunden till OAI**

Under senare delen av 1900-talet har nya kommunikationsmodeller vuxit fram, som alternativ till de vetenskapliga tidskrifterna. Genom faxen och datorer fick forskarna nya sätt att kommunicera på, t.ex. via e-mail (Tomaiuolo & Packer 2000). En något senare metod för att kommunicera vetenskapliga resultat är s.k. e-prints. E-prints har sin föregångare i pre-prints. Ett pre-print kan definieras på lite olika sätt. Dess huvudsakliga betydelse är vetenskaplig artikel som har blivit granskad och accepterad av en vetenskaplig tidskrift, men ännu inte publicerats (Tomaiuolo & Packer 2000). Även artiklar som inte har blivit granskade eller accepterade kallas ibland pre-prints. Motsvarande granskade och publicerade artiklar kallas post-prints (Pinfield et al. 2002).

Ett e-print definieras ibland som ett pre-print som gjorts tillgängligt i elektronisk form på webben. Men ett e-print ses ibland som något mer än bara en elektronisk version av ett pre-print. Så definierar t.ex. the American Physical Society e-prints som "alla elektroniska arbeten som av författaren gjorts tillgängliga utanför den traditionella publiceringsvärlden" (Tomaiuolo & Packer 2000; *min översättning*). Det finns även olika ståndpunkter när det gäller vad som bör publiceras i ett e-printarkiv. Enligt Pinfield, Gardner och MacColl (2002) finns det inom vissa discipliner ett motstånd mot att tillgängliggöra pre-prints för allmänheten, medan det inom andra discipliner finns en etablerad pre-print kultur. Med e-printarkiven kan man dock se ett begynnande motstånd mot traditionella sätt att publicera nya vetenskapliga rön och ett spirande försök till arkiverings- och publiceringsalternativ, som innebär besparingar i både tid och kostnader. Ofta framhålls också den enkelhet med vilken resultat kan publiceras på

webben, allt som behövs är ett program som kan skapa HTML-dokument och en web-server (Tomaiuolo & Packer 2000).

E-prints blev snabbt en uppskattad kommunikationsform inom vissa forskarkollektiv (Lagoze & Van de Sompel 2001). En starkt bidragande orsak var att det nya mediet gav forskarna möjlighet att mycket snabbt sprida sina alster till olika kolleger. I den traditionella publiceringskedjan kan det ta upp till ett år innan en färdig artikel kommer i tryck. Inom de s.k. hårda vetenskaperna, t.ex. fysik, där utvecklingen går mycket snabbt, är detta extra bekymmersamt (Lagoze & Van de Sompel 2001). Tomaiuolo och Packer (2000) påpekar att det var inom fysik och matematik som forskare först fick tillgång till högpresterande datorer och att e-prints gav möjlighet till inkludering av grafik, en nog så viktig faktor. Det har under senare år talats mycket om ytterligare en förklaring till den omstrukturering vi nu bevittnar och det är de galopperande prisökningarna för prenumerationer på vetenskapliga tidskrifter (se t.ex. Rabow 2001 och Hagerlid 2002). En del håller dock för mindre troligt att det är just prisökningarna som fått allt fler forskare att välja vägar utanför den traditionella publiceringskedjan (Hagerlid 2002, s.98). Det är nog snarare så att prisökningarna i stället är en av bibliotekens starkaste drivkrafter när det gäller att bistå forskarna på vägen mot ”öppna arkiv” och andra alternativ.

Ganska omgående började man inom vissa forskarkollektiv bygga upp en typ av server-baserade arkiv för dessa e-prints. Det första e-printarkivet skapades 1991 av Paul Ginsparg, en fysiker vid Los Alamos National Laboratory i USA. Arkivet fick namnet *arXiv.org e-Print archive* (arXiv.org ePrint archive 2003). Även om e-printarkiv är vanligast inom naturvetenskap och medicin, har det skapats arkiv även inom samhällsvetenskap och humaniora. Ett av de mer etablerade är *Cogprints* (Cogprints 2003), som innehåller e-prints inom bl.a. psykologi, antropologi och filosofi. Ett annat relativt nyetablerat arkiv är *E-LIS* (E-LIS 2003), ett e-printarkiv för forskare inom Biblioteks- och informationsvetenskap. Till en början var e-printarkiven mer av ett komplement till än en ersättning för traditionell publicering. Inom vissa vetenskaper, t.ex. fysik, är idag elektronisk egenarkivering en absolut nödvändighet och publicering i traditionella tidskrifter har blivit mer av en formalitet (Tomaiuolo & Packer 2000).

4.4. Open Archives Initiative

Allteftersom nya e-printarkiv etablerades uppstod behovet av att kunna göra samsökningar i arkiven. I oktober 1999 arrangerades ett möte i Santa Fe, New Mexico, utifrån "en tro på att interoperabilitet mellan arkiven var nyckeln till att öka dessa arkivs inflytande" (Lagoze & Van de Sompel 2001; *min översättning*). Förutom att öka arkivens inflytande trodde man också att en sammanslagning skulle leda till mer utbyte av kunskap mellan discipliner, något som ansågs som positivt utifrån det faktum att allt fler discipliner arbetar tvärvetenskapligt. Vid mötet deltog forskare, dataspecialister och bibliotekarier. Målet var att skapa de tekniska och organisatoriska förutsättningarna för samsökning av världens alla olika e-printarkiv, oavsett inom vilka discipliner arkiven hör hemma. Deltagarna vid mötet lyckades ganska omgående att komma överens om en teknisk lösning, medan de organisatoriska överenskommelserna till en början var av mer informell karaktär. I den tekniska överenskommelsen, kallad *the Santa Fe Convention*, ingick bl.a. en gemensam metadatastandard och en metod för insamling av metadata, uppbyggd kring ett protokoll kallat *the OAI Protocol for Metadata Harvesting* (Lagoze & Van de Sompel 2001).

Under år 2000 fortsatte arbetet inom OAI oförtrutet. Initiativtagarna blev varse att det fanns ett stort intresse för OAI:s arbete långt utanför e-printkollektivet. Framförallt visade det sig finnas ett starkt intresse bland forskningsbiblioteken i USA. Denna insikt ledde till att vissa förändringar gällande den tekniska specifikationen genomfördes, för att protokollet skulle passa även andra användargrupper. Man insåg också under denna tid att man, för att öka initiativets trovärdighet, måste bygga upp en organisatorisk infrastruktur. En ledningsgrupp tillsattes och följande uttalande om OAI:s mål kungjordes:

The Open Archives Initiative develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content. The Open Archives Initiative has its roots in an effort to enhance access to e-print archives as a means of increasing the availability of scholarly communication. Continued support of this work remains a cornerstone of the Open Archives program. The fundamental technological framework and standards that are developing to support this work are, however, independent of both the type of content offered and the economic mechanisms surrounding that content, and promise to have much broader relevance in opening up access to a range of digital materials. As a result, the Open Archives Initiative is currently an organization and an effort explicitly in transition, and is committed to exploring and enabling this new and broader range of applications. As we gain greater knowledge of the scope of applicability of the underlying technology and standards being developed, and begin to understand the structure and the culture of the various adopter communities, we expect that we will have to make continued evolutionary changes to both the mission and organization of the Open Archives Initiative. (Lagoze & Van De Sompel 2001)

Uttalandet visar, anser jag, på flera intressanta aspekter av OAI:s arbete. Kopplingen till vetenskaplig kommunikation och e-printarkiv är fortfarande stark, även om man säger sig vilja bidra till ökad tillgänglighet för alla typer av digitalt material. Utveckling av tekniker och standarder betonas, men man antyder också att kunskap om strukturer inom och kulturella skillnader mellan olika användargrupper är viktig i det fortsatta arbetet. Likaså visas en öppenhet gentemot framtida förändringar av både initiativets inriktning och tekniska lösningar.

Det första officiella protokollet för samsökning i olika arkiv, *the OAI Protocol for Metadata Harvesting 1.0* (OAI-PMH 1.0), släpptes i januari 2001 och startskottet gick för en ett år lång försöksperiod. Fler och fler e-printarkiv började använda sig av OAI-PMH 1.0. Till en början var arkiven huvudsakligen disciplinorienterade, men allteftersom gjordes även försök till institutionsbaserade e-printarkiv (Rusch-Feja 2002). I en artikel i Ariadne argumenterar Pinfield, Gardner och MacColl (2002) för att utvecklingen mot fler institutionsbaserade arkiv bör uppmuntras. De menar att institutioner, t.ex. universitetsbibliotek, har de ekonomiska, tekniska och organisatoriska resurser som behövs för att understödja utvecklingen av e-printarkiv. Utöver detta har institutionerna ett egenintresse av att upprätta arkiv, nämligen att stärka sin ställning i forskarvärlden. I juni 2002 kom en andra version av protokollet, OAI-PMH v.2.0. Efter

månader av experimenterande, tester och revideringar ansåg man sig redo att släppa en uppdaterad och stabil version.

Innan jag går över till att beskriva OAI: s återvinningsmodell ska jag nämna några ord om själva initiativets namn, som till en början möttes av protester från vissa håll. Det är främst termen *archive* som förorsakat diskussion. Ett arkiv har för vissa yrkesgrupper en mycket bestämd innebörd, som kan sägas innefatta både själva den institution ett arkiv utgör och dess syfte, ett långsiktigt bevarande av särskilt betydelsefulla dokument. Den innebörd OAI ger termen har sitt ursprung inom eprint-kollektivet, där man med arkiv menar en samling vetenskapliga, digitala dokument. OAI definierar *archive* på följande sätt: ”en behållare (*repository*) för lagrad information” (Lagoze & Van De Sompel 2001; *min översättning*). Termen *open* definierar initiativtagarna utifrån ett ”arkitektoniskt perspektiv”, vilket innebär att OAI syftar till att öka tillgängligheten till information av olika slag. *Open* ”är inte det samma som ’gratis’ eller ’obegränsad access’ till de arkiv som har valt att följa OAI: s modell (Lagoze & Van De Sompel 2001; *min översättning*).

I synnerhet har tvetydigheten kring termen *archive* bidragit till att den engelska termen *repository* under senare tid har börjat användas i stället (Pinfield 2003). Lunds universitet använder en försvenskad form av termen, *repositor*. Förutom i avsnitt 4.5, där jag bl.a. redogör för terminologin inom OAI, kommer jag inte att använda termen *repositor* (eller *repository*) utan hålla mig till ”öppna arkiv” alternativt ”arkiv”, väl medveten om termens dubbeltydighet och att *repositor* mycket väl kan bli den term som i framtiden kommer att användas i svenska sammanhang.

4.5. OAI Protocol for Metadata Harvesting 2.0

I följande avsnitt kommer jag att förklara vissa grundläggande begrepp inom OAI: s återvinningsmodell.

4.5.1. Metadata

En av grundbultarna i OAI: s återvinningsmodell är metadata. Metadata är numera en allmänt accepterad term inom Biblioteks- och informationsvetenskap, men kan ändå kräva sin förklaring. Ibland definieras metadata helt kort som ”data om data”. Utifrån en sådan definition kan t.ex. en recension av en bok vara metadata (Deegan & Tanner 2002, s. 113). Enligt en något mer snäv definition förknippas metadata ofta med digitala resurser och blir då en modern term för ett gammalt välkänt begrepp inom biblioteksvärlden, nämligen bibliografisk information. D.v.s. information om eller en beskrivning av ett dokument vars syfte är att identifiera och underlätta organisation, sökning, urval, åtkomst och utbyte av dokument. Det kan vara titel, författarnamn, publiceringsår, etc. Denna typ av metadata kan även kallas *deskriptiv metadata* (Deegan & Tanner 2002, s. 116ff). Deegan och Tanner nämner även *strukturell metadata* och *administrativ metadata*. Strukturell metadata underlättar för användaren att se vilken typ av dokument han eller hon har med att göra. Strukturerna i en barnbok och i en vetenskaplig artikel är t.ex. mycket olika. Strukturerna i ett dokument underlättar även för användaren att ”navigera” sig fram i dokumentet (Deegan & Tanner 2000, s. 118). Strukturell metadata är extra viktig i digitala resurser. Deegan och Tanner gör jämförelsen mellan en vanlig bok och ett digitalt objekt. Det är lätt att se när man har läst $\frac{3}{4}$ av en pappersbok. I en e-bok däremot måste detta visas på något annat sätt.

Administrativ metadata innehåller uppgifter om t.ex. filformat och uppgifter som har att göra med ekonomiska och legala aspekter (Deegan & Tanner 2002, s. 116ff). I den här uppsatsen kommer jag att ge metadata en tämligen vid definition och avser med metadata *digitala beskrivningar av digitala resurser*. Med digitala resurser avser jag text, bild, ljud och andra typer av digital information.

Som det ser ut idag skapas metadata i en mängd olika format. Björkhem och Lindholm (2001, s. 6) menar att det stora antalet olika metadataformat beror på att det dels är olika typer av resurser som beskrivs med hjälp av metadata, dels att metadata i sig används för olika syften. Formatens komplexitet varierar från det enklaste tänkbara till mycket komplexa. Dempsey och Heery (1997) har delat in olika metadataformat i tre kategorier - enkla, strukturerade och "rika" format. Den typ av metadata som sökmaskiner på webben skapar, är ett exempel på ett enkelt format. Dublin Core kan räknas som ett strukturerat format och MARC¹ som ett rikt format (Deegan & Tanner 2002, s. 121). MARC utvecklades för maskinläsbara bibliografiska beskrivningar av tryckt material, för användning i bibliotekskataloger on-line, s.k. OPAC, men används idag även för beskrivningar av digitala resurser. Dublin Core växte fram som en lösning på den omöjliga situation biblioteken ställdes inför, när antalet digitala resurser formligen exploderade under 1990-talet. MARC-formatet ansågs av vissa som allt för komplext och tidskrävande för det stora antal digitala resurser biblioteken hade att katalogisera. Initiativtagarna till Dublin Core ville skapa ett format anpassat just för digitala resurser, som till sin utformning skulle vara enklare och därmed billigare att använda än MARC-formatet. Lösningen på problemet blev ett enkelt formulär med vars hjälp författaren, eller någon annan, kan skapa metadata enligt ett standardiserat format. Metadata fogas sedan in som en slags katalogpost i inledningen av det digitala dokumentet.

Utöver mer välkända format som Dublin Core och MARC finns det ett antal mindre kända och oftast mer komplexa metadataformat, som har vuxit fram ur olika användargrupperns särskilda behov. Några olika exempel är TEI², EAD³ och LOM⁴.

4.5.2. Interoperabilitet

I takt med att antalet digitala samlingar ökar blir kravet på interoperabilitet allt större, d.v.s. ett utbyte ska kunna ske mellan olika samlingar. Detta behov av interoperabilitet kan finnas såväl inom ett digitalt bibliotek, som mellan olika digitala bibliotek eller andra typer av informationssystem. Det finns dock många hinder på vägen mot en väl utvecklad interoperabilitet (Deegan & Tanner 2002, s.140f). Två av dessa hinder har med teknik och metadata att göra. För att interoperabilitet överhuvudtaget ska vara möjlig, måste samlingarna bygga på någon slags struktur, d.v.s. vara uppbyggda som någon typ av databas. Flera olika tekniker kan dock användas för att bygga databaser, t.ex. Microsoft Access eller MySQL, vilket kan utgöra ett hinder för interoperabilitet. Olika metadataformat kan också utgöra ett hinder för utbyte databaser emellan. Lunds universitet kartlade under 2001 ett 50-tal av sina runt 100 digitala samlingar. Bland detta begränsade antal samlingar återfanns inte mindre än fem olika databastekniker och 6-7 olika metadataformat, vilket säger en del om problemets komplexitet (Kjellberg 2001).

¹ MARC = Machine Readable Cataloguing

² TEI = Text Encoding Initiative

³ EAD = Encoded Archival Description

⁴ LOM = Learning Object Metadata

När det gäller problemet med inkompatibla metadataformat, vore ett gemensamt format för *alla* digitala resurser den optimala lösningen ur återvinnings synpunkt. Många när en dröm om detta, fast snarare utifrån ett "meta-metadataformat" som tillåter sökningar mellan olika metadataformat (Koehler 2002). Ett av de mer omfattande försöken att skapa ett meta-metadataformat är Tim Berners-Lees och hans kollegers arbete med "the Semantic web". Den semantiska webben bygger på flera nya tekniker och standarder, bl.a. XML⁵ och RDF⁶, där RDF är det ramverk som ska möjliggöra interoperabilitet mellan olika metadataformat och i förlängningen leda till en slags katalogisering av webben (Berners-Lee, Hendler & Lassila 2001).

För att återgå till de digitala biblioteken, så har man även där skapat verktyg för att underlätta interoperabilitet mellan olika metadataformat. Z39.50, ett av de första meta-metadataformaten, används flitigt inom biblioteksvärlden. Z39.50 bygger på en "komplex uppsättning kommunikationsprotokoll" (Breeding 2002a; *min översättning*). Ett protokoll är "en uppsättning regler som underlättar för datorer att utbyta information över nätverk" (Deegan & Tanner 2002, s.149; *min översättning*) och medger därmed samsökning i befintliga databaser. Exempel på andra viktiga protokoll i nätverkssammanhang är IP⁷ och HTTP⁸. Det finns många fördelar med komplexa protokoll som Z39.50, t.ex. att protokollet medger access till information i realtid och på så sätt ger uppdaterad information. Komplexa, funktionella protokoll medför dock ofta implementeringssvårigheter och höga kostnader (Armes, se Lagoze & Van De Sompel 2001).

Enligt Lagoze och Van De Sompel (2001) finns det när man talar om metadata, två välkända krav att ta hänsyn till och som kan vara svåra att tillmötesgå samtidigt. Det handlar dels om tidigare nämnda interoperabilitet mellan format, dels om olika användargrupperns specifika behov, som kan behöva uttryckas i mer komplexa metadataformat. När det gäller interoperabilitet kan man, menar Lagoze och Van De Sompel (2001), välja att lägga bördan på dem som ska använda metadata, t.ex. de digitala biblioteken. Dessa får då lösa problemet med heterogena metadataformat med hjälp av protokoll som Z39.50. Rakt motsatt väg skulle vara ett standardformat med de eventuella begränsningar det skulle innebära för olika användarspecifika behov. OAI har valt en tredje väg, där man hoppas kunna förena de två kraven ovan. Enligt upphovsmännen ska OAI:s protokoll inte ses som en ersättning för, utan ett komplement till mer komplexa lösningar som Z39.50 (Lagoze & Van De Sompel 2001).

Vad OAI hoppas kunna åstadkomma med OAI-PMH 2.0 är ett lågröskelprotokoll, som genom sin utformning ska vara enkelt och billigt att installera. OAI gör en tydlig uppdelning mellan "simple resource discovery" och "community-specific description" (Lagoze & Van De Sompel 2001). Hur dessa två funktioner i praktiken ska kunna samspela inom ett och samma protokoll återkommer jag till i avsnitt 4.5.3.

Utifrån den återvinningsmodell som OAI förespråkar kan man urskilja tre typer av aktörer. Jag kommer nedan att förklara de olika aktörernas roller. Det finns ännu inte någon gängse svensk terminologi för de begrepp som används inom OAI. I den mån svenska översättningar används, är det i de flesta fall mina egna och den engelska

⁵ XML=Extensible Markup Language

⁶ RFD=Resource Description Framework

⁷ IP= The Internet Protocol

⁸ HTTP=The Hypertext Transfer Protocol

termen följer då inom parentes. Ord i kursiverad stil är grundläggande begrepp inom OAI:s återvinningsmodell

Till att börja med har vi de aktörer som tillhandhåller data, d.v.s. tillhandahåller resurser (*resources*). Resurserna ska vara försedda med metadata. Huruvida resurserna ska vara digitala eller inte, tar OAI inte ställning till (Lagoze et al. 2002). Med tanke på bakgrunden till hela projektet håller jag för troligt att det huvudsakligen kommer att handla om just digitala resurser. Dessa resurser kan vara dokument i en forskare privata arkiv, i ett ämnesbaserat arkiv eller i ett institutionsarkiv. Denna typ av aktör kallas dataleverantör (*dataproducer*).

De olika resursernas metadata ska finnas lagrad i en s.k. repositör (*repository*). Denna repositör kan även innehålla själva resurserna, men dessa kan också finnas lagrade i en annan databas. Varje resurs motsvaras i repositören av en *item*, "en slags behållare (*container*) som lagrar eller dynamiskt genererar metadata..." (Lagoze et al. 2002; *min översättning*). Eftersom en och samma resurs kan beskrivas med hjälp av olika metadataformat, Dublin Core plus ett eller flera användarspecifika format, säger man att en *item* kan ha olika poster (*records*), en post för varje metadataformat. Varje *item* har en identifikator (*identifier*) som är unik inom respektive repositör. Metadatan för de olika resurserna måste vara skapad enligt av OAI fastställda regler. En repositör måste även ha ett program installerat som kan processa de sex OAI-frågor (*requests*) som protokollet är uppbyggt kring. I maj 2003 finns 83 repositörer registrerade på OAI: s hemsida (Open Archives Initiative 2003). Exempel på repositörer är tidigare nämnda *arXiv.org e-Print archive*, *Cogprints* och *E-LIS*.

Nästa typ av aktör är de som kallas informationstjänster (*serviceproviders*). Med hjälp av OAI-PMH samlar informationstjänsten in metadata från utvalda repositörer, från vilka beror på informationstjänsternas olika målgrupper. Därefter organiseras den insamlade metadatan i en ny databas. Via ett gränssnitt kan användaren sedan söka i informationstjänstens databas, bland de i förväg insamlade metadataposterna. I posterna bör det finnas en länk till själva resurserna oavsett var dessa finns lagrade. Fördelen med denna typ av informationstjänst är att användarna inte själva behöver hålla reda på och söka i alla olika arkiv som kan vara av intresse. I stället använder man sig av den informationstjänst som bäst passar de egna behoven. Hur själva informationstjänsten utformas kommer att skifta utifrån tänkt användargrupp. Tanken är att informationstjänsten, i förhållande till de arkiv de hämtar metadata ifrån, ska innehålla ett antal tilläggstjänster i mer eller mindre sofistikerad form. Exempel på sådana tilläggstjänster kan t.ex. vara olika browse- eller sökfunktioner, att dokument kan rankas utifrån en användarprofil eller en bedömning av olika dokumenters impact factor⁹ (se t.ex. Dobratz & Matthaei 2003). I maj 2003 finns 12 informationstjänster registrerade på OAI: s hemsida. Ett exempel på en informationstjänst är *my.OAI* (*my.OAI* 2003), som tillåter sökning i 15 OAI-kompatibla databaser. Ett annat exempel är *perseus.tufts.edu* (*perseus.tufts.edu* 2003), där användaren kan hitta texter, bilder, m.m. från olika historiska epoker. Söktjänsten *perseus.tufts.edu* innehar också rollen av dataleverantör.

⁹ Impact Factor (IF)= "Ett mått på hur ofta en genomsnittlig artikel i en viss tidskrift citerats under ett givet år. IF beräknas som relationen mellan citeringar och publicerade citeringsbara artiklar" (Rodman 2003).

Den tredje aktören är användarna. Informationstjänsternas användare kommer troligtvis till stor del vara forskare, d.v.s. de personer, institutioner eller universitet som också agerar som dataleverantörer, även om arkivens ”öppenheten” ska gälla också andra intresserade.

I en undersökning genomförd av Open Archives Forum (Dobratz & Matthaei 2003) kring utvecklingen av OAI-aktiviteter i Europa, konstaterar man att det är fler dataleverantörer än informationstjänster som har implementerat OAI-tekniken. Man förutspår dock att antalet informationstjänster kommer att öka stadigt framöver. Undersökningen visar också att ett flertal dataleverantörer även fungerar som informationstjänster eller har för avsikt att utvecklas därefter. En förklaring till att antalet dataleverantörer är fler än antalet informationstjänster kan vara att det är förhållandevis enkelt att göra redan existerande digitala samlingar OAI-kompatibla, under det att det krävs betydligt större insatser för att skapa en OAI-söktjänst (Pinfield et al. 2002; Breeding 2002a).

4.5.3. Metadata enligt OAI:s modell

Själva kärnan i OAI:s samsökningsprotokoll är, som tidigare nämnts, användningen av ett gemensamt metadataformat. Till en början användes ett format som kallades *the Open Archives Metadata Element Set*. Detta format ansågs ganska omgående vara allt för starkt knutet till utbyte av information mellan e-printarkiv. För att bredda protokollets presumtiva användningsområde valde man att utgå från det redan existerande och tämligen spridda formatet Dublin Core.

Metadataformatet Dublin Core skapades ursprungligen vid ett möte 1995, i staden Dublin, Ohio, där ett antal bibliotekarier, datatekniker och nätverksspecialister deltog. Bibliotekens behov av ett gemensamt format för beskrivning av elektroniska resurser var vid denna tid mycket stort. Deltagarna hade följande mål med det nya formatet: Det skulle vara enkelt att använda, lätt att förstå, internationellt och utbyggbart. Vid mötet enades man om en uppsättning metadataelement, till en början 13 och senare 15 till antalet. Dessa 15 element kallas *The Dublin Core Metadata Element Set*, med vars hjälp författaren själv eller någon annan kan beskriva ett dokument. Till denna basala del av Dublin Core hör *Title, Creator, Subject, Description, Publisher, Contributor, Date, Format, Type, Identifier, Source, Language, Relation, Coverage* och *Rights*. Alla element är frivilliga, kan upprepas och ordningen på dem bestämmer den som gör beskrivningen. De 15 grundelementen kan preciseras genom s.k. *qualifiers*. När det gäller användningen av *qualifiers* är Dublin Core tämligen tillåtande, vilket gör att behoven inom olika användargrupper eller metadataprojekt kan styra utformningen, dock inom vissa givna gränser (Degerstedt 2002).

Med metadata enligt Dublin Core vill OAI skapa något av en minsta gemensamma nämnare för de olika arkiven. Enkelhet är ledordet i denna del av OAI-PMH. Trots att det med hjälp av Dublin Cores *qualifiers* finns stora möjligheter att göra specifika resursbeskrivningar, rekommenderar OAI att endast de 15 grundelementen används. Inget av de 15 elementen är obligatoriska inom OAI:s modell. Metadata enligt denna modell blir, enligt Lagoze och Van De Somple (2001), tämligen grov och tillåter enbart "simple resource discovery". OAI uppmuntrar dock olika användargrupper att utveckla egna metadatapraktiker (Lagoze & Van De Somple 2001). För att ge olika användargrupper möjlighet till mer specifika resursbeskrivningar, utifrån de egna

behoven, kan den obligatoriska metadatan enligt Dublin Core kompletteras med en eller flera parallella metadataposter, det som Lagoze & Van De Sompel kallar "community specific descriptions". När det gäller dessa parallella poster finns inga begränsningar, förutom att de måste vara strukturerade i enlighet med uppmärkningspråket XML. Exempel på disciplinspecifika metadataformat som används av OAI-kompatibla dataleverantörer är MARC, TEI, EAD och IMS¹⁰ (Breeding 2002b).

4.5.4. Metadatapost enligt OAI: s modell

Varje resurs, d.v.s. varje dokument i en repositor motsvaras av en *item*. Beskrivningen av en *item* kan, som vi har sett, uttryckas i flera olika metadataformat, en post för varje format. En post identifieras genom en *item*-identifikator, i kombination med ett metadataprefix som talar om vilket format det rör sig om och datum för postens tillblivelse. Posten är uppdelad i tre delar:

- *header*: Innehåller information som är identisk för alla poster, d.v.s. information oberoende av vilket metadataformat posten är skapad i. Informationen i huvudet består av två komponenter, en identifikator och en datumstämpel som indikerar förändringar i metadatan (när metadatan är skapad, ändrad eller borttagen);
- *metadata*: Det måste finnas minst *en* metadatapost och den ska vara formaterad enligt Dublin Core. Dublin Core-posten kan kompletteras med poster i andra användarspecifika metadataformat;
- *about*: Denna del är frivillig. Ska i de fall den används innehålla information om metadatan, t.ex. ursprung och eventuella restriktioner gällande användning.

Det är viktigt att tillägga att identifikatorn för en *item* inte är lika med identifikatorn för resursen i fråga. Hur resursernas identifikatorer utformas ligger utanför OAI:s ansvarsområde, men kan t.ex. utgöras av en URL-adress¹¹. OAI förespråkar att en länk upprättas mellan posten/posterna och resursen i fråga, med hjälp av resursens identifikator, som kan anges i Dublin Core-elementet *Identifier* (Lagoze & Van de Sompel 2001; Breeding 2002b).

På nästa sida finns ett exempel på en post, hämtad från Lagoze et al. (2002). Posten är en beskrivning av ett dokument arkiverat i *arXiv.org E-print archive*. Jag har markerat de tre grunddelarna, *header*, *metadata* och *about* med fetstil. Det är från poster liknande denna som söktjänsterna, med hjälp av protokollets sex frågor, hämtar information. I Dublin Core-elementet *dc: identifier* finns en länk till dokumentet ifråga.

¹⁰ IMS = Instructional Management System

¹¹ URL = Uniform Resource Locator

```

<header>
  <identifier>oai:arXiv.org:cs/0112017</identifier>
  <timestamp>2002-02-28</timestamp>
  <setSpec>cs</setSpec>
  <setSpec>math</setSpec>
</header>
<metadata>
  <oai_dc:dc>
    xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
    http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
      <dc:title>Using Structural Metadata to Localize Experience of Digital
      Content</dc:title>
      <dc:creator>Dushay, Naomi</dc:creator>
      <dc:subject>Digital Libraries</dc:subject>
      <dc:description>With the increasing technical sophistication of both information
      consumers and providers, there is increasing demand for more meaningful
      experiences of digital information. We present a framework that separates
      digital object experience, or rendering, from digital object storage and manipulation,
      so the rendering can be tailored to particular communities of users.
      </dc:description>
      <dc:description>Comment: 23 pages including 2 appendices,
      8 figures</dc:description>
      <dc:date>2001-12-14</dc:date>
      <dc:type>e-print</dc:type>
      <dc:identifier>http://arXiv.org/abs/cs/0112017</dc:identifier>
    </oai_dc:dc>
</metadata>
<about>
  <provenance
    xmlns="http://www.openarchives.org/OAI/2.0/provenance"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/provenance
    http://www.openarchives.org/OAI/2.0./provenance.xsd">
    <originDescription harvestDate="2002-02-02T14.10.02Z" altered="true">
    <baseURL>http://the.oa.org</baseURL>
    <identifier>oai:r2.org:klik001</identifier>
    <timestamp>2002-01-01</timestamp>
    <metadataNamespace>http://openarchives.org/OAI/2.0/oai_dc/
    </metadataNamespace>
    </originDescription>
  </provenance>
</about>

```

4.5.5. Protokollfrågor

Det protokoll som OAI har valt att bygga sin modell på är mycket enkelt utformat och består av endast sex frågor (Lagoze et al. 2002). Med hjälp av dessa sex frågor samlar de olika informationstjänsterna in information från olika repositorer. En *OAI-PMH harvester* är analog med olika sökmaskiners robotar, som systematiskt återvinner och indexerar HTML-dokument, men med den stora skillnaden att en *OAI-harvester* endast kan samla in metadata från OAI-kompatibla dataleverantörer. Både informationstjänsternas frågor och dataleverantörernas svar överförs via webbens http-protokoll. En konsekvens av OAI-protokollets enkelhet är att det inte går att göra urval av metadataposter, t.ex. utifrån ämne eller författare. Tanken är i stället att metadataposter ska överföras "en masse" till olika informationstjänster, som sedan i sin tur utformar ett gränssnitt med tilläggstjänster. OAI-protokollet tillåter dock två typer av urval, utifrån datum och set (en dataleverantör kan bygga upp sin samling i olika set).

De frågor protokollet bygger på är följande:

1. *GetRecord* - återvinner en enskild metadata-post. För att detta ska kunna ske måste identifikatorn och önskat metadataformat anges.
2. *Identify* - återvinner information om en repositör.
3. *ListIdentifiers* - återvinner en lista över identifikatorerna för posterna i en repositör.
4. *ListMetadataFormat* - återvinner en lista över vilka format en repositör använder sig av.
5. *ListRecords* - återvinner poster från en repositör. Urval möjligt utifrån set och datum.
6. *ListSets* - återvinner information om vilka eventuella set en repositör har valt dela in sitt material i.

Detta var OAI:s modell i mycket korta drag. För den som är intresserad finns protokollet i sin helhet på OAI:s webbplats (Lagoze et al. 2002). I följande avsnitt kommer jag att diskutera några problem förknippade med *the OAI Protocol for Metadata Harvesting*, som under senare tid uppmärksammats från brittiskt håll.

4.5.6. Kritiska röster

En del av de undersökningar som hitintills genomförts när det gäller OAI och OAI-PMH handlar till stor del om inventeringar av olika OAI-aktiviteter - hur många dataleverantörer respektive informationstjänster som finns, vilka olika hård- och mjukvaror som används, metadataformat etc. (Dobratz & Matthaei 2003; Pinfield, Gardner & MacColl 2002; Pinfield 2003). Kritikerna är enligt min bedömning fortfarande få. Det kan delvis bero på att OAI:s återvinningsmodell är så pass ny som den är. Dessutom är den del i ett långt mycket större systemskifte, där gamla och nya kommunikationssätt lär leva sida vid sida under lång tid framöver. Enligt vissa personer kommer stora delar av den formella vetenskapliga kommunikationen i framtiden att ske i digital form. Om detta ska bli möjligt är det, enligt min mening, ett gigantiskt projekt forskarvärlden och andra intressenter har att genomföra.

Pinfield, Gardner och MacColl (2002) presenterar några praktiska frågor som institutioner har att ta ställning till vid upprättandet av "öppna arkiv". Artikelns bygger på två brittiska universitets erfarenheter av att skapa eprint-arkiv. Pinfield, Gardner och MacColl är överlag positiva till "öppna arkiv" men lyfter fram flera viktiga frågeställningar rörande bl.a. arkivens innehåll, metadata och hur arkiven ska fyllas med innehåll.

OAI-protokollet ger de tekniska förutsättningar för interoperabilitet mellan olika metadataformat, men säger ingenting om arkivens innehåll. Pinfield, Gardner och MacColl (2002) menar att det måste utformas en policy för de samlingar som ska ingå i arkiven. Dels behöver man ta ställning till vilken typ av dokument som ska accepteras, t.ex. huruvida både pre- och post-prints ska accepteras. De som inrättar arkiv måste också ta ställning till vilken typ av dokumentformat som ska tillåtas. Pinfield, Gardner och MacColl är t.ex. tveksamma till om HTML-dokument ska tillåtas. HTML uppfattas av Pinfield, Gardner och MacColl som ett allt för flytande uppmärkningspråk som är svårt att validera.

Pinfield, Gardner och MacColl (2002) nämner också arkiveringsmomentet som en källa till svårigheter. Trots att det program som används vid de två berörda universiteten har en funktion för att föra över dokument till arkiven, uppstod problem vid själva arkiveringen, p.g.a. både ointresse och bristande kunskaper hos forskarna. Pinfield, Gardner och MacColl förespråkar att t.ex. biblioteket vid ett universitet fungerar som en mellanhand och har ansvar för själva arkiveringen.

En mycket viktig fråga som Pinfield, Gardner och MacColl (2002) uppmärksammar är hur man ska få forskare att vilja arkivera sina alster. Att upprätta ett arkiv är en sak, att få forskare att vilja använda arkiven en helt annan, konstaterar de. Det handlar då om att både få forskare att *arkivera* material och att *använda* material som andra har arkiverat. Detta är ett uppbyggnadsarbete som måste ske i två faser. I den första fasen upprättas en "demonstrator". I denna "demonstrator" bör man till en början arkivera material som redan finns tillgängliga, t.ex. på olika institutioners webbsidor. I fas nummer två gäller det att nå en kritisk massa av innehåll. För att nå detta mål måste forskare övertygas om fördelarna med "öppna arkiv", som enligt Pinfield, Gardner och MacColl är många. Inom fysik finns det en etablerad pre-printkultur. Så är inte fallet inom många andra discipliner. Tvärtom, hävdar Pinfield, Gardner och MacColl, finns det ett utbrett motstånd inom många andra discipliner mot ett allmänt tillgängliggörande av pre-prints. För att attrahera även dessa forskare bör inte e-printarkiv framställas som en ersättning för publicering i traditionella tidskrifter, utan som ett alternativ till traditionell publicering. Det skulle då huvudsakligen handla om att arkivera post-prints. Skillnaderna mellan olika discipliner är så stora att vissa universitet t.o.m. bör överväga att ha flera olika arkiv, t.ex. ett för pre-prints och ett för post-prints. Pinfield, Gardner och MacColl understryker att de ansvariga för e-printarkiven i utsträckning måste ta hänsyn till olika forskarsgruppers behov och sätt att arbeta.

Pinfield, Gardner och MacColl (2002) påpekar också hur viktigt det är att metadatan skapas på rätt sätt. OAI har valt en mycket enkel form av metadata. Trots det, eller kanske på grund av det, kan det uppstå problem med kvaliteten på metadatan, hävdar de. Detta problem hänger ihop med att det är författarna som ska skapa metadatan. En experimentell söktjänst, ARC Service, har rapporterat ett antal problem som har att göra med inkonsekvent metadatagenerering, alltifrån "olika varianter av stavning eller

dataformat till mer komplexa problem som olika ämnesdeskriptorer” (Pinfield, Gardner & MacColl 2002; *min översättning*). ARC Service menar också, enligt Pinfield, Gardner och MacColl, att det är svårt att normalisera ”denna typ av metadata för att t.ex. skapa ett index” (2002; *min översättning*). Pinfield, Gardner och MacColl hävdar att någon typ av metadatagranskning innan dokumenten offentliggörs är mycket viktig. De anför ytterligare en svaghet med anknytning till metadata enligt OAI: s modell och det är att konventionella sökmotorer inte plockar upp metadata enligt Dublin Core. Pinfield, Gardner och MacColl (2002; *min översättning*) förespråkar därför en vidareutveckling av mjukvaror som kan ”översätta OAI-kompatibel metadata till sökmotorsanpassad data”.

Pinfield (2003) vidareutvecklar tankarna i Pinfield, Gardner och MacColl (2002) och redogör i mer systematiserad form för erfarenheterna av ”öppna arkiv” i Storbritannien under senare år. Pinfield menar att vi nu befinner oss vid en kritisk punkt i utvecklingen av institutionella öppna arkiv och att ”de två till tre kommande åren kan komma att bli avgörande för utvecklingen på lång sikt” (Pinfield 2003; *min översättning*). Från att ha varit en fråga för ”ett relativt litet antal entusiaster” börjar ”öppna arkiv” bli något som intresserar allt fler inom informationsrelaterade verksamheter och universitet (Pinfield 2003; *min översättning*). Pinfield sammanfattar de brittiska erfarenheterna i fyra punkter. För det första är det framtida arbetet med ”öppna arkiv” *inte* en teknisk fråga. Den teknik som hitintills har utvecklats fungerar tillfredsställande, det gäller både själva protokollet och olika mjukvaruprogram. Vad som däremot är viktigt att diskutera framöver är frågor relaterade till ”collection management”, t.ex. vilka dokumenttyper och dokumentformat som ska tillåtas, vem ska skapa metadata, metadatakvalitet och enligt vilken standard metadata ska skapas, bevarandenaspekter, m.m. Ekonomiska frågor är också viktigt att diskutera, t.ex. framtida kostnader för underhåll av olika arkiv. Den absolut största utmaningen är dock, enligt Pinfield, att fylla arkiven med innehåll. Egenarkivering kräver en ”kulturell förändring bland forskare” (Pinfield 2003; *min översättning*) som endast kan komma till stånd genom upplysning om arkivens fördelar. Det krävs också en ökad medvetenhet bland dem som propagerar för ”öppna arkiv”, om de barriärer som hindrar forskare från att arkivera i och använda sig av ”öppna arkiv”. I detta arbete är det också viktigt att ta hänsyn till olika disciplin-specifika frågor.

*eprints.org*¹² (2003) har också uppmärksammat problemet med uppmärkning och olika dokumentformat. Med uppmärkning menas att man med hjälp av t.ex. HTML-taggar märker upp olika funktionella delar av ett dokument - titel, rubriker, tabeller, stycken m.m., för att bl.a. underlätta informationsåtervinning. *eprints.org* (2003) konstaterar att uppmärkning har kommit att bli allt viktigare när det gäller digitala dokument. Många dokument skapas idag i HTML, PDF, eller Word, format som enligt *eprint.org* varken är tillräckligt rika eller flexibla för att klara framtida krav på digitala dokument, det gäller såväl sofistikerade funktioner som bevarandenaspekter. Ett mer avancerat format under utveckling är XML, ett format som anses ha en stabilitet och utvecklingsmöjligheter långt utöver t.ex. HTML. Att märka upp dokument med XML är dock en mycket kostbar procedur. Flera av de kommersiella förlagen använder idag XML i sina elektroniska tidskrifter. Så länge ”öppna arkiv” är mer av ett komplement till, än en

¹² *eprints.org*, med hemvist vid universitetet i Southampton, England, ingår i *Open Citation Project*, ett *DLI2 International Digital Libraries Project*. Projektet finansieras av *Joint Information System Committee*, (*JISC*) vid *Higher Education Funding Councils* och *National Science Foundation*.

ersättning för, publicering i kommersiella tidskrifter, är bristen på XML-uppmärkning i arkiverade dokument inget större problem. I framtiden kan utvecklingen dock, enligt *eprint.org*, ta två vägar. Antingen förblir situationen som den är idag, d.v.s. ”öppna arkiv” kommer huvudsakligen att innehålla kopior av XML-uppmärkta artiklar publicerade i traditionella tidskrifter. Förlagens XML-uppmärkta dokument och de fördelar de för med sig kommer att vara förbehållna dem som har råd att betala. Övriga får nöja sig med arkivens dokument i HTML, PDF, etc., med de begränsningar det för med sig. Ett annat tänkbart scenario är att arkiven kommer att bli så populära att de kommersiella förlagen, på grund av vikande efterfrågan, inte längre kan bekosta dyr XML-uppmärkning. Arkiven skulle då bli tvungna att gå över till XML, för att kunna möta framtida användares krav och behov. *eprints.org* tror att det kommer att utvecklas användarvänliga XML-editorer, vilka skulle kunna bidra till att lösa problemet med de höga kostnaderna för XML-uppmärkning.

För att XML-uppmärkning ska fungera måste dokumenten i sig vara välstrukturerade. En vetenskaplig artikel är ofta uppbyggd enligt en fastställd struktur, vilket underlättar en eventuell uppmärkning i XML. Denna struktur kan dock variera mellan olika ämnesområden (Rodman 2003). Betydligt svårare är det, menar jag, att märka upp t.ex. magisteruppsatser i Biblioteks- och informationsvetenskap, där strukturen skiftar från uppsats till uppsats. Vid Blekinge Tekniska Högskola gjordes för ett tag sedan ett försök till uppmärkning i XML av svenska examensarbeten, men försöket fick avbrytas just p.g.a. bristen på en gemensam struktur i uppsatserna. I stället har Blekinge Tekniska Högskola, tillsammans med fyra andra svenska universitet och högskolor, på försök byggt upp en söktjänst för examensarbeten, kallat *Arkiv Ex – Sverige* (Arkiv Ex-Sverige 2003), som bygger på metadata-repositorier och ett gemensamt metadataformat, allt enligt OAI:s modell (Lundberg & Lindholm 2003).

4.6. Sammanfattning avsnitt 4

”Öppna arkiv” bygger på en kommunikationsmodell som utvecklades bland fysiker i början av 1990-talet. I kommunikationsmodellen ingår ett utbyte av vetenskapliga artiklar i elektronisk form och en typ av fulltextdatadaser/arkiv, i vilken de elektroniska dokumenten lagras. Sammanslutningen Open Archives Initiative bildades 1999. Målet var att hitta de tekniska och organisatoriska lösningarna för samsökning av alla de olika eprint-arkiv som hade skapats under 1990-talet, främst inom naturvetenskapliga discipliner. Initiativtagarnas förhoppning var att en möjlighet till samsökning av arkiven skulle ökas deras inflytande (Lagoze & Van de Sompel 2001).

OAI:s modell för samsökning av ”öppna arkiv” bygger på metadataformatet Dublin Core och ett protokoll kallat *the OAI Protocol for Metadata Harvesting* (OAI-PMH). I juli 2002 släpptes en andra, uppgraderad version av OAI-PMH. Protokollet bygger på endast sex frågor och enkelheten sägs bidra till både användarvänlighet och låga kostnader. Dublin Core valdes för att det är ett förhållandevis känt format och enkelt att använda. Med Dublin Core kan endast grova beskrivningar göras. För att möta olika användargrupper mer specifika kan behov, kan beskrivningen i Dublin Core kompletteras med beskrivningar i andra format (Lagoze & Van de Sompel 2001).

Inom OAI:s arkiverings/publiceringsmodell finns tre aktörer - dataleverantörer, informationstjänster och användare. Dataleverantörer kallas de som genom privata arkiv, ämnesarkiv eller institutionsarkiv levererar elektroniska resurser försedda med

metadata. Informationstjänsterna samlar in metadata från de olika dataleverantörerna och bygger upp söktjänster anpassade efter den tilltänkta målgruppens behov (Lagoze et al. 2002).

Trots att OAI-PMH utvecklades under relativt kort tid, fungerar den teknik som protokollet bygger på bra och den framtida utvecklingen av denna nya modell för vetenskaplig kommunikation anses inte vara en teknisk fråga. Däremot har andra svårigheter visat sig efterhand. Nedan följer i punktform några olika faktorer som har visat sig skapa problem när några universitet i Storbritannien under senare år byggt upp institutionsarkiv (Pinfield, Gardner & MacColl 2002; Pinfield 2003; eprints.org 2003):

- Metadata: OAI har valt ett format som sägs vara enkelt att använda. Trots detta har det uppstått problem när det gäller kvaliteten på metadadatan. Vem ska skapa metadata? Vilka metadataformat ska godkännas?
- Dokumenttyper: Vilka dokumenttyper ska arkiveras? (Pre-prints, post-prints, monografier, etc.)
- Formattyper: Vilka formattyper ska godkännas? (HTML, PDF, XML etc.)
- Bevarandaspekter: Hur ska bevarandefrågan lösas på lång sikt?
- Ekonomiska aspekter på lång sikt. Arkiven är billiga att bygga upp, men kan på lång sikt kräva desto större resurser.
- Hur ska man få forskare att vilja använda arkiven?

De problem som sammanfattas ovan kommer alla att behöva finna sin lösning, för att "öppna arkiv" ska bli en effektiv och säker modell på lång sikt. Vissa av problemen har större anknytning till Biblioteks- och informationsvetenskap än andra. Inom detta område finns t.ex. stora kunskaper om metadata, som kan och bör användas i uppbyggnaden av "öppna arkiv". Fortsatt forskning kring metadata är också viktig. Jag anser att även forskning kring metadata som inte direkt anknyter till "öppna arkiv", kan ha relevans för det fortsatta arbetet inom Open Archives Initiative. Så tror jag t.ex. att det projekt som pågår vid School of Information and Library Science i North Carolina, Metadata Generation Research Project, kommer att generera kunskap som kan vara värdefull vid uppbyggnaden av "öppna arkiv". Flera av ovanstående problem måste också lösas genom ett samarbete mellan olika professioner, t.ex. informationsvetare och dataspecialister. De ekonomiska aspekterna måste också diskuteras. Hur mycket ska ett systemskifte få lov att kosta och vad får det för konsekvenser för andra verksamheter vid t.ex. ett universitetsbibliotek? Sammanfattningsvis kan jag konstatera att Open Archives Initiative är ett spännande initiativ med stor potential. Allt fler frågor har dock börjat ställas under senare tid och initiativets arbete med återvinningsmodellens tekniska och kunskapsorganisatoriska delar kommer med all sannolikhet att behöva utvecklas i framtiden. OAI-PMH 2.0 lär inte vara den sista versionen vi får se.

5. Vägen mot ”öppna arkiv”

I följande avsnitt kommer jag att ta mig an uppsatsen andra syfte, att undersöka ”öppna arkiv” utifrån ett historiskt, socialt och kulturellt perspektiv, i synnerhet i förhållande till de ”öppna arkivens” presumtiva användare. Jag kommer här att med hjälp av tidigare genomförda undersökningar och olika texter om alternativa publiceringsmodeller försöka förstå och förklara det motstånd mot ”öppna arkiv” som sägs existera bland många forskare.

5.1. En ny modell, men för vem?

Andersen (2002) har visat hur nya kommunikationstekniker genom tiderna har förändrat människans sätt att ”producera, kommunicera, lagra, återvinna och organisera kunskap” (s.37; *min översättning*). Datorer och Internet har onekligen förändrat forskares sätt att producera och kommunicera kunskap, en förändring som på intet sätt är avslutad. Enligt Kling och McKim (2000, s.1306) är nya kommunikationsmönster en av de största förändringarna inom forskarvärlden av idag. Hur omstruktureringen av den vetenskapliga kommunikationen kommer att sluta, är det idag ingen som vet. Det går dock att urskilja några olika sätt att se på framtiden. Vissa hävdar med emfas att elektroniska medier kommer att dominera inom vetenskaplig kommunikation, medan andra förhåller sig mer kritiska och ifrågasättande.

Några av de mest ivriga förespråkarna för en övergång till elektroniska medier är *Paul Ginsparg*, grundare av det tidigare nämnda arkivet för fysik, *Steven Harnad*, redaktör för den elektroniska tidskriften *Psychology* och *Andrew Odlyzko*, verksam vid Digital Technology Center vid universitetet i Minnesota. Det finns naturligtvis många fler, men dessa tre förekommer ofta i debatten. Enligt Ginsparg är frågan snarare när, och inte om, den dag kommer då all spridning av vetenskaplig forskningslitteratur kommer att ske i elektronisk form (se Kling & McKim 2000, s.1310). Även inom andra sammanslutningar, som SPARC och BOAI, finns det en förhoppning om att vi verkligen står inför ett paradigmskifte när det gäller vetenskaplig kommunikation. Det är inte helt lätt att fastslå huruvida förespråkarna menar vetenskaplig kommunikation inom *alla* discipliner eller inom *alla naturvetenskapliga* discipliner. Enligt Kling & McKim uttrycker åtminstone vissa entusiaster att ”en enda modell för elektronisk vetenskaplig kommunikation passar alla discipliner” (2000, s.1310; *min översättning*).

Svårigheten att förstå huruvida den nya modellen verkligen är tänkt att gälla alla discipliner kan, enligt min mening, bero på användningen av ordet vetenskap. I engelsktalande länder skiljer man mellan *science* och *scholarship*, där man med *science* menar naturvetenskapliga discipliner och med *scholarship* i synnerhet humanistisk lärdom. På svenska kan dock *science* även betyda vetenskap i största allmänhet (scholarship 1980; science 1980). Ibland används även termen *natural sciences* när man avser naturvetenskap. När det gäller den engelska termen *social sciences* är inte heller den oproblematiserad. En del ifrågasätter huruvida discipliner inom *social sciences* verkligen kan kallas *sciences* och hävdar att denna term endast bör användas för naturvetenskapliga discipliner. För en mer omfattande diskussion kring svårigheterna med att definiera olika termer och forskningsområden och om olika syn på vad som bör kallas *science*, se Meadows (1998, s.39ff). I olika artiklar om det nya kommunikationsparadigmet förekommer alla varianter: *scholarly publishing*, *scientific*

research literature, scholarly communication, scientists, scholars, etc., vilket ändå tyder på att förespråkarna menar att de alternativ som nu börjar dyka upp, ska kunna användas inom alla discipliner.

Det är även svårt, anser jag, att fastställa huruvida omstruktureringen ska gälla alla dokumenttyper som används inom det traditionella publiceringsparadigmet: tidskriftsartiklar, monografier, konferenspapers, rapporter, avhandlingar, etc. Någon säger uttryckligen att "uppropet" endast gäller granskade forskningsartiklar och pre-prints (Suber 2003), medan andra menar att ett institutionsarkiv kan innehålla allt från pre-prints till granskade tidskriftartiklar, monografier och läromedel (Johnson 2002).

Det tycks dock vara så att olika förespråkare för alternativa arkiverings- och publiceringsmodeller överlag anser att dessa passar det stora flertalet av såväl samhällsvetenskapliga som humanistiska och naturvetenskapliga discipliner, även om de senare uttryckligen nämns mer frekvent. När det gäller vilka dokumenttyper som anses lämpliga att arkivera/publicera i "öppna arkiv" talas det, så vitt jag kan bedöma, betydligt oftare om pre-prints och granskade tidskriftsartiklar, än andra typer av dokument.

Även om en positiv inställning dominerar debatten, finns det de som ifrågasätter om den elektroniska modellen är lika lämplig för alla discipliner. Kling & McKim (2000) menar att övergången till elektroniska medier är långt mer komplex än vad förespråkarna, av Kling & McKim kallat "The Electronic Publishing Reform Movement", hävdar (2000, s.1309). Ett annat exempel är Julie M. Hurd (2000, s.1279ff). Hon tror att olika discipliner kommer att välja olika vägar i transformeringen av den vetenskapliga kommunikationen.

Många av de debattörer som under senare år diskuterat en omstrukturering av den vetenskapliga kommunikationen tycks vara överens om att det hos ett (oroväckande) stort antal forskare finns ett motstånd mot alternativ till traditionell vetenskaplig publicering. Ett citat av Savenije & Smith får illustrera denna ståndpunkt:

Several initiatives have been started to realise other ways of scientific publishing, using modern information technology. In several disciplines, however, a relatively large number of scientists are still reluctant to make use of completely new ways of publishing.
(Savenije & Smith 2002, s.1)

Flera andra skribenter resonerar på liknande sätt, se t.ex. Tomaiuolo & Packer (2000), Harnad (2001), Crow (2002), Pinfield, Gardner och MacColl (2002) och Pinfield (2003). Även om antalet "öppna arkiv" stadigt växer, även inom humaniora och samhällsvetenskap, tycks det finnas en del barriärer på vägen mot att fylla arkiven med den kritiska massa som är avgörande för om arkiven ska få den framträdande roll som många hoppas på. Dessa barriärer kan troligtvis vara av olika karaktär. Även om Pinfield (2003) menar att den teknik som finns är tillförlitlig, kan det troligtvis finnas barriärer av teknisk karaktär. Det är också tänkbart att det finns ekonomiska barriärer, barriärer med anknytning till legala frågor och rent disciplinrelaterade barriärer, utifrån kulturella, historiska och sociala skillnader.

I avsnitt 5.3 har jag för avsikt att redogöra för tänkbara barriärer som kan förhindra, försvåra eller fördröja en övergång till ”öppna arkiv” och andra alternativa modeller. Jag har också för avsikt att försöka undersöka om det finns några tecken på att motståndet är större bland forskare inom humaniora och samhällsvetenskap, än bland naturvetenskapligt inriktade forskare. Jag kommer i stort att behandla humaniora och samhällsvetenskap som en homogen grupp discipliner och kontrastera dessa mot naturvetenskapliga discipliner. Jag är medveten om att detta är en grov förenkling av verkligheten. Det finns stora skillnader mellan humaniora och samhällsvetenskap. Meadows (1998) och Lindholm-Romantschuk (1998) har visat på några av dem. Dels är flertalet discipliner inom samhällsvetenskap betydligt yngre än de humanistiska disciplinerna. Filosofi t.ex. kan vi härleda till antikens Grekland, medan sociologi uppstod som vetenskaplig disciplin i början av 1900-talet. En annan viktig skillnad är val av metod. Discipliner inom humaniora använder sig oftare av tolkande metoder, till skillnad från samhällsvetenskaperna som ”tenderar att använda metoder inlånade från naturvetenskaperna” (Lindholm-Romantschuk 1998, s.103; *min översättning*). Likaså finns det ett antal tvärvetenskapliga discipliner, inte minst Biblioteks- och informationsvetenskap, som hämtar inspiration, teori och metoder från flera olika ämnesområden. Att jag huvudsakligen talar om naturvetenskap, kontra samhällsvetenskap och humaniora, får ses som ett praktiskt grepp för att göra min studie något mer hanterbar. Utom i något enstaka fall kommer jag inte heller att beröra ingenjörsvetenskaperna.

För att nå en ökad förståelse för orsakerna till motståndet mot egenarkivering och för att försöka ta reda på om det förekommer några skillnader mellan olika discipliner, ämnar jag använda mig av tidigare genomförda användarstudier och studier av hur forskare inom humaniora och samhällsvetenskap kommunicerar med varandra. Jag kommer även att referera till en studie av hur forskare inom några av dessa discipliner i allmänhet förhåller sig till ny teknik, artiklar som berör ämnet samt en konferens om vetenskaplig kommunikation inom humaniora och samhällsvetenskap.

Ett sätt att ta reda på hur forskare förhåller sig till ”öppna arkiv” skulle vara att undersöka existerande ämnesarkiv och institutionsarkiv. När det gäller ämnesarkiv skulle man kunna studera hur antalet arkiv fördelar sig mellan olika discipliner och i vilken utsträckning forskare bidrar med material till de olika arkiven. När det gäller institutionsarkiv skulle man kunna studera i vilken grad olika discipliner bidrar med material till arkivet. Några sådana undersökningar har mig veterligen inte genomförts.

Ett annat sätt skulle vara att genom intervjuer eller enkätundersökningar undersöka forskares attityder till egenarkivering. Några sådana studier har också genomförts under senare år. Enligt min bedömning finns det dock tämligen begränsat med forskning kring forskares inställning till ”öppna arkiv”. Jag kommer att hänvisa till två studier med anknytning till temat egenarkivering. Den första är en användarstudie av Hunt (2002), där hon jämför inställningen till egenarkivering bland användare, respektive ickeanvändare, av två olika arkiv. Den andra undersökningen (Gadd, Oppenheim & Probetz 2003) pågår för närvarande i Storbritannien, inom något som kallas RoMEO-projektet. Där avser man att genomföra en mer omfattande undersökning kring frågor rörande ”öppna arkiv” och intellektuell upphovsrätt. Gadd, Oppenheim & Probetz (2003) hänvisar till ytterligare en studie från 2002 (Swan), som jag dock inte har kunnat få tillgång till.

När det gäller olika discipliners inställning till ”öppna arkiv” och andra alternativ, antyder vissa debattörer att motståndet är större bland forskare inom humaniora och samhällsvetenskap än inom naturvetenskap. The Knight Higher Education Collaborative (2002, s. 209; *min översättning*) menar att forskare inom humaniora och samhällsvetenskap har varit långsammare med att se ”digital publicering som den huvudsakliga kanalen för spridning” av vetenskapliga resultat. Hunt (2002; *min översättning*) skriver att arkiv ”fortfarande används av en minoritet av författare (30-40 %) och bara inom ett litet antal fält (huvudsakligen fysik, matematik och datavetenskap)”. Tomaiuolo & Packer (2000) konstaterar att de allra flesta e-printarkiv innehåller naturvetenskaplig information, från discipliner såsom ”astronomi, kemi, datavetenskap, matematik och fysik”. Vissa områden inom humaniora och samhällsvetenskap har ”nyligen har hakat på trenden” men ”ligger dock tydligt efter när det gäller antalet servrar” (Tomaiuolo & Packer 2000; *min översättning*). Det kan tyckas vanskligt att hänvisa till så pass gamla uttalanden inom ett område där utvecklingen går mycket snabbt. Jag håller dock för troligt att den eftersläpning, som enligt Tomaiuolo & Packer fanns inom humaniora och samhällsvetenskap år 2000, har fortsatt även efter övergången från e-printarkiv till ”öppna arkiv”.

5.2. Den historiska och sociala kontextens betydelse

Innan jag går över till att undersöka barriärer mot egenarkivering kommer jag att redogöra för två texter som sätter in vetenskaplig kommunikation i ett socialt och historiskt sammanhang. Den ena texten är en monografi skriven av Jack Meadows, en forskare som under lång tid studerat vetenskaplig kommunikation, främst via tryckta medier. Den andra texten är en artikel av Rob Kling och Geoffrey McKim. De fokuserar, i högre grad än Meadows, på elektroniska medier.

Jack Meadows arbetar som professor vid institutionen för Information and Library studies, vid Loughborough University i Storbritannien. Meadows har även erfarenhet från bl.a. astronomi och vetenskapshistoria. Enligt White & McCains studie (1998, s.334) är Meadows huvudsakliga verksamhetsområden citeringsanalys, vetenskaplig kommunikation och citeringsteori.

Meadows har studerat vetenskaplig kommunikation under flera årtionden, huvudsakligen utifrån ett processperspektiv, d.v.s. varför och hur forskare kommunicerar med varandra (Borgman 2000, s.414). 1974 gav Meadows ut boken *Communication in Science*, enligt Borgman (2000, s. 412) en milstolpe inom forskning kring vetenskaplig kommunikation. I *Communication in Science* diskuterar Meadows hur forskare inom naturvetenskap kommunicerar med varandra (Borgman 2000, s. 413). 24 år senare publicerade Meadows boken *Communicating Research*, i vilken han vidgar sin horisont och gör jämförelser mellan olika discipliner inom naturvetenskap, humaniora och samhällsvetenskap. Boken publicerades 1998, d.v.s. året innan OAI drogs igång och Meadows skriver därför, av naturliga skäl, inte mycket om ”öppna arkiv” och andra alternativ till traditionell vetenskaplig publicering. Han nämner *arXiv.org e-Print archive*, men endast med några få rader (Meadows 1998, s. 149; s. 203). Meadows utgångspunkt är att forskning och vetenskaplig kommunikation står inför stora förändringar, i och med övergången till elektroniska medier. Ett av flera temata i boken *Communicating research* är att kunskap om hur forskare förhåller sig till det tryckta mediet ”kan ge kunskap om hur forskarkollektivet kommer att reagera på en övergång till ett elektroniskt kommunikationsmedium” (Meadows 1998, s. 1-2; *min*

översättning). Att en övergång till ett nytt medium kommer att påverka den vetenskapliga kommunikationen tycks Meadows inte tvivla på. På så vis finns det en likhet mellan Meadows och Andersen (2002). Meadows går dock ett steg längre än Andersen, genom att hävda att tekniska förändringar av olika slag inte nödvändigtvis får samma effekt inom alla discipliner. I förordet till *Communicating research* skriver han följande rader:

Researchers in these two fields [astronomi och vetenskapshistoria, min anm.] see the world from different angles; their ideas on the nature of acceptable knowledge differ, and their research communities are organised differently. Consequently, their handling of information is dissimilar. Changes affecting the world of research as a whole do not necessarily have identical impacts on research communication in the sciences and on that in the humanities. (Meadows 1998, s. ix)

I *Communicating Research* gör Meadows en mycket omfattande genomgång av vitt skilda områden inom forskning och vetenskaplig kommunikation. Han börjar med en beskrivning av vetenskapens framväxt och hur uppdelningen mellan naturvetenskap, samhällsvetenskap och humaniora tog form. Därefter reflekterar han kring vem som forskar, varför och med vilka resultat. Meadows redogör även för olika kommunikationskanaler och procedurer vid vetenskaplig kommunikation. Meadows menar att det finns stora skillnader mellan olika forskningsområden och då inte bara utifrån vad man forskar om, utan också hur man ser på bl.a. kunskap och relationen mellan teori och praktik och vilka metoder man föredrar att använda. Meadows poängterar att gränserna mellan de tre olika forskningsområdena inte är knivskarpa, men att nämnda skillnader kan få konsekvenser för hur forskare inom olika områden förhåller sig till dominerande kommunikationstekniker och till hur de väljer att kommunicera sina resultat (1998, s. 39ff).

Enligt Meadows är det två faktorer som styr på vilket sätt forskare väljer att kommunicera. Dessa två faktorer är dels det *medium* som används för överföring av information, dels olika forskarkollektivs *behov*, utifrån både egenskaperna av producenter och mottagare av information. Meadows hävdar att hur forskare väljer att använda tryckta medier, i större utsträckning styrs av olika behov än av tekniska förändringar (1998, s. 1). Han menar också att även om det är de enskilda forskarna som i sista hand avgör om eller hur ett medium ska användas, påverkas de, ofta omedvetet, av olika normer med historiskt och socialt ursprung inom det egna forskarkollektivet. Detta kollektiva minne påverkar hur forskare under en övergång från ett medium till ett annat kommer att presentera och använda information menar Meadows (1998, s. 239).

Med hjälp av ett otal olika studier visar Meadows att det finns faktiska skillnader mellan olika discipliner när det gäller vetenskaplig kommunikation, alltifrån antal författare per tidskriftsartikel, till vilken dokumenttyp olika forskare föredrar att citera, till hur många tidskriftsartiklar respektive monografier forskare inom olika discipliner gett upphov till. Meadows diskuterar dessa frågor huvudsakligen utifrån olika discipliners förhållande till tryckta medier, men gör även en del reflektioner utifrån övergången till elektroniska medier.

Meadows bok ger en något splittrad bild av dessa uppenbarligen skiftande kommunikationsmönster inom olika discipliner. Det beror troligtvis på den spännvidd

som finns i de undersökningar han bygger på och det är svårt att på ett adekvat sätt redogöra för hela bokens innehåll. Meadows visar dock på ett tydligt sätt att det finns skillnader mellan olika forskningsområden, både vad gäller forskningen i sig och hur den kommuniceras. Meadows text kan, enligt min åsikt, bidra till en ökad förståelse för olika forskares attityder gentemot ”öppna arkiv”. Likaså kan en av Meadows slutsatser ha relevans för min frågeställning, nämligen den att det snarare är olika forskares behov, än kommunikationsmediets möjligheter, som avgör hur olika forskare väljer att närma sig ett nytt medium. Det går inte, utifrån Meadows text, att förutspå hur forskare inom olika discipliner i slutändan kommer att förhålla sig till elektroniska medier, men jag anser att man åtminstone kan dra slutsatsen att alla forskare inte kommer att förhålla sig på exakt samma sätt.

Kling och McKim, verksamma vid Center for Social Informatics vid Indiana University School of Library and Information Science, verkar i samma anda som Meadows. De menar att forskarvärlden är långt ifrån så homogen som vissa debattörer framhåller. De anser att det finns stora skillnader mellan hur forskare inom olika discipliner kommunicerar med varandra och att dessa skillnader med all sannolikhet kommer att bestå under lång tid framöver (Kling & McKim 2000). De argumenterar för sin sak utifrån ett ”social shaping perspective”, ett teoretiskt perspektiv som de har hämtat från MacKenzie & Wajcman (se Kling & McKim 2000, s. 1311). Ett perspektiv enligt denna teori, kallad *Social shaping of technology*,

... focuses on the ongoing dynamic between a technology and a community, as the technology is developed, used, shaped, reconfigured, and reconstituted within the community. (Kling & McKim 2000, s. 1311)

Innebörden i detta perspektiv är att ett antal olika sociala krafter styr i vilken mån och på vilka sätt ny teknik kommer att anammas av samhällets olika aktörer. Kling och McKim hämtar även en del begrepp från ett annat teoretiskt perspektiv, *New Institutionalism*, med rötter inom bl.a. sociologi, ekonomi och studier av hur organisationer fungerar. Enligt detta perspektiv styrs olika aktörer inom en organisation av bl.a. rutiner och vanor som fastställts och förstärks av ”centraliserade institutioner” (Kling & McKim, 2000, s. 1311; *min översättning*). Enligt Kling och McKim är institutioner en av de sociala krafter som styr människors beteende och då snarare som ”en källa till tröghet än radikal förändring” (2000, s. 1311; *min översättning*). Kling och McKim sätter sitt eget perspektiv i kontrast till de perspektiv som enbart ser till olika mediers tekniska särdrag och/eller hävdar att utvecklingen mot elektroniska medier är mer eller mindre självgående och ofrånkomlig (2000, s. 1311).

Enligt min åsikt är ett perspektiv i Meadows och Kling och McKims anda att föredra, framför perspektiv som enbart utgår från tekniken och dess inneboende möjligheter. Utifrån teorin att *samspelet* mellan människans behov och teknikens möjligheter avgör vilken betydelse ny teknik får, kommer jag i följande avsnitt att diskutera olika barriärer som kan tänkas bidra till vissa forskargrupper avaktande inställning till egenarkivering.

5.3. Om attityder till och tänkbara barriärer mot egenarkivering

En publiceringsform måste kunna garantera vissa grundläggande saker för att attrahera forskare. Kling och McKim (1999, s. 897f) nämner tre faktorer som har betydelse för huruvida en publiceringsform kan anses vara effektiv eller ej. Dessa tre faktorer är den valda publiceringsformens förmåga att garantera publicitet, tillförlitlighet och tillgänglighet. Vilken av dessa tre faktorer som kan anses viktigast finns det troligtvis skilda åsikter om. Halliday (2001), som för övrigt bygger sin undersökning på bl.a. Kling och McKim (1999), kommer fram till att det som är mest avgörande för huruvida en publiceringsform kan anses effektiv eller ej, är graden av tillgänglighet. Detta fann jag något förvånande. Jag skulle nog ha gissat att tillförlitlighet är mer betydelsefullt. Bakom Kling och McKims modell döljer sig dock en rad olika omständigheter som jag ska försöka redogöra för nedan.

Jag har i litteraturen kunnat urskilja ett antal faktorer som sägs ha betydelse för forskares inställning till alternativa modeller inom vetenskaplig kommunikation. Nedan följer en lista över olika faktorer som nämns. En del faktorer förekommer betydligt mer frekvent än andra. Jag har dock inte rangordnat de olika faktorerna.

- oklarheter kring kvalitetskontroll vid egenpublicering
- oklarheter kring upphovsrätt vid egenpublicering/rädsla för plagering
- ekonomiska faktorer
- tekniska faktorer
- bevarandenaspekter
- bruket av pre-prints inom en disciplin
- behovet av snabb spridning av forskningsresultat
- huruvida det finns ett aktivt "invisible college"
- hur allmänt förekommande storskaliga projekt är
- geografisk spridning av forskningsteam
- utbyte av stora dataset vid tvärvetenskapligt samarbete
- vilken roll patent spelar vid skyddande av "intellectual property"
- antalet kärntidskrifter inom en disciplin

De fem första punkterna har utkristalliserat sig under arbetets gång. De åtta sista punkterna är hämtade från Hurd (2000, s. 1283). Jag kan se ytterligare en faktor, som inte uttryckligen nämns i litteraturen, men som skulle kunna vara en tänkbar orsak till att utvecklingen av "öppna arkiv" går långsammare inom vissa discipliner. Denna faktor är

- vilken publiceringsform som är vanligast förekommande inom en disciplin

Jag kommer nedan att diskutera en del av dessa fjorton faktorer, några mer, andra mindre, för att se hur de eventuellt skulle kunna påverka inställningen till egenarkivering. Jag har valt att dela in följande avsnitt i underrubriker. Flera av faktorerna påverkar dock varandra och ibland återfinns flera faktorer under en och samma rubrik. Likaså återkommer vissa faktorer under flera olika rubriker. I anslutning till varje faktor kommer jag att föra en diskussion om vilken betydelse den kan tänkas ha för humanistiskt och samhällsvetenskapligt inriktade forskares inställning till egenarkivering.

5.3.1. Bevarandenaspekter

I den litteratur jag har läst talas det förvånansvärt lite om bevarandenaspekter vid elektronisk publicering. Det kan naturligtvis bero på att jag huvudsakligen valt litteratur som närmar sig elektronisk publicering utifrån ett användarperspektiv. Jag kommer inte heller att gå närmare in på bevarandefrågan. Hur ett långsiktigt bevarande av digitala resurser ska kunna garanteras är huvudsakligen en fråga för dem som utvecklar systemet, även om jag håller för troligt att bevarandefrågan säkerligen påverkar forskares inställning till egenarkivering. Hur stort utrymme bevarandenaspekter kommer att få framöver beror troligtvis på om ”öppna arkiv” blir ett medium för enbart informell kommunikation eller för både informell och formell kommunikation. Om arkivering/publicering i alternativa modeller i framtiden ersätter traditionell publicering, kommer frågan om hur och vilka av dessa dokument som ska bevaras för eftervärlden bli mycket viktig. Begreppet arkivering inom OAI står huvudsakligen för ”deponering” av artiklar. Inom ett projekt kallat the Open Archival Information System (OAIS) arbetar man sedan en tid tillbaka utifrån en mer traditionell definition av arkivering, d.v.s. hur dokument ska kunna bevaras för eftervärlden (Pinfield, Gardner & MacColl 2002).

5.3.2. Kvalitetskontroll

En förmodad brist på kvalitetskontroll genom peer review vid egenarkivering, är ett förhållande som ganska ofta nämns som en tänkbar orsak till vissa forskares motstånd. Den kvalitetskontroll som sker innan artiklar publiceras i olika tidskrifter är en viktig del av traditionell vetenskaplig publicering. Vissa ifrågasätter om denna kvalitetskontroll kan bibehållas när vetenskapligt material i stället arkiveras i ”öppna arkiv” eller publiceras i ”open access journals” (se t.ex. Rowland 1997). Andra menar att kvalitetskontroll mycket väl går att lösa även inom alternativa modeller (se t.ex. Harnad 2001; Odlyzko 2002). Hur avgörande denna fråga i förlängningen blir beror troligtvis på vilken typ av material som kommer att arkiveras i de olika arkiven. I de fall man i enbart arkiverar kopior av redan granskade artiklar kan detta inte vara något problem. I de fall enbart pre-prints arkiveras, i avvaktan på publicering i en tidskrift med peer review, bör detta inte heller vara något större problem. När arkivering däremot ska *ersätta* publicering i granskade tidskrifter, kan frågan om kvalitetskontroll komma att få stor betydelse.

Kling och McKim (2000, s. 1306ff) är några av alla dem som diskuterar hur synen på peer review påverkar olika forskares inställning till ”öppna arkiv”. De har studerat i vilken utsträckning elektroniska medier används inom tre olika discipliner, teoretisk partikelfysik, molekylärbiologi och informationssystem. De analyserar i sin artikel de olika forskargruppernas inställning till egenarkivering och i vilken mån forskarna är villiga att dela med sig av, och ta del av andras, icke kvalitetsgranskat material. Kling och McKim fastslår att det finns skillnader mellan de tre disciplinerna. Inom teoretisk partikelfysik sker en stor del av den formella kommunikationen via arkiv, även om publicering i tidskrifter fortfarande fyller vissa funktioner relaterade till långsiktigt bevarande, prestige och anslag. Inom molekylärbiologi finns det ett litet antal servrar, men den formella kommunikationen sker huvudsakligen via tidskrifter. Forskare inom informationssystem slutligen har utvecklat IS-World, en webbaserad samling av bl.a. länkar och artiklar som dock inte går att jämföra med de arkiv forskare inom fysik använder sig av. Kling och McKim tycker sig kunna urskilja fyra omständigheter som

kan påverka forskares attityder till egenarkivering. Den första omständigheten är kostnaden för forskningsprojektet. Dyra projekt skulle kunna ge en tendens till ökat samarbete, via t.ex. ”öppna arkiv”. Den andra omständigheten är i vilken utsträckning forskare är medvetna om vad andra inom samma fält håller på med. Här hävdar de att forskare inom fysik är mer medvetna om vad andra inom samma fält sysslar med än forskare inom t.ex. sociologi. Kling och McKim menar att om öppenheten utåt är stor, kan rädslan för att dela med sig av icke granskat material bli mindre. Den tredje omständigheten som Kling och McKim nämner, har att göra med att patent och affärshemligheter kan bidra till en större försiktighet när det gäller att dela med sig av material. Till sist tror Kling och McKim att det inom discipliner med ett begränsat antal kärntidskrifter, t.ex. astrofysik, inte finns något behov av att öka tillgängligheten genom nya kanaler. De avslutar sin artikel med att dra slutsatsen att uppdelningen kommer att bestå mellan de discipliner där traditionell peer review har minskat i betydelse och de där traditionell peer review kommer att förbli betydelsefull.

Kling och McKim berör i sin artikel flera av de ovan nämnda faktorerna: projektkostnader, storskalighet, grad av medvetenhet om andra forskares arbete, patentfrågor, antalet kärntidskrifter inom en disciplin, bruket av pre-prints, geografisk spridning av forskningsteam och utbyte av stora dataset. De menar att teoretisk partikelfysik skiljer sig från de andra två disciplinerna. Inom fysik finns det, till skillnad från de flesta andra discipliner, en lång tradition av att dela med sig av sina resultat via pre-prints. Detta utbyte skedde redan innan webben utvecklades för drygt 10 år sedan. Storskaliga projekt är också vanligt förekommande inom fysik, likaså den geografiska spridningen av forskningsteam och utbyte av stora dataset. Jag håller för troligt att det är färre projekt inom humaniora och samhällsvetenskap, än inom naturvetenskap, som kan karaktäriseras som storskaliga och kostsamma, även om det naturligtvis existerar undantag, t.ex. inom arkeologi. Detta antagande gör jag utifrån det faktum att det inom humaniora och samhällsvetenskap finns mindre ekonomiskt utrymme än inom naturvetenskap, vilket bl.a. The Knight Higher Education Collaborative (2002, s. 209) diskuterar. Detta skulle, enligt Kling och McKims resonemang, kunna leda till en större slutenhet inom humaniora och samhällsvetenskap. Man kan naturligtvis ifrågasätta om dessa forskningsområden verkligen är mer slutna. Likaså kan man diskutera Kling och McKims åsikt att sociologer skulle vara mindre medvetna om kollegernas arbete än t.ex. fysiker. Som vi ska se längre fram anser en del, även bland humanister och samhällsvetare själva, att det finns en syn på dessa två discipliner som två något isolerade och defensiva forskningsområden. En sådan självbild skulle naturligtvis kunna bidra till en slutenhet gentemot andra forskare.

Jag antar att patentfrågor knappast spelar någon roll inom humaniora och samhällsvetenskap. När det gäller antalet kärntidskrifter har jag inte haft möjlighet att undersöka hur det ser ut inom humaniora och samhällsvetenskap och kan inte ge mig på att spekulera hur denna faktor eventuellt påverkar forskare inom humaniora och samhällsvetenskap.

Utifrån alla de artiklar jag har läst råder det ingen tvekan om att peer review och tillit är frågor som kommer att spela stor roll för den fortsatta utvecklingen av ”öppna arkiv”. Oklarheter kring peer review är med all sannolikhet en av orsakerna till att inte alla forskare lika snabbt anammat idén om egenarkivering. Tyvärr säger Kling och McKims artikel inte så mycket om humanister och samhällsvetares inställning till peer review. Men jag skulle kunna tänka mig att behovet av traditionell peer review är större

inom tvärvetenskapliga discipliner och enligt Lindholm-Romatschuk (1998, s. 31) är humanistiska och samhällsvetenskapliga discipliner tvärvetenskapliga i större utsträckning än naturvetenskapliga discipliner. Antagandet om ett större behov av traditionell peer review inom tvärvetenskapliga discipliner, gör jag utifrån det faktum att forskare inom tvärvetenskapliga discipliner förmodas hämta information från områden de kanske inte känner till lika bra som sitt eget och behovet av granskade artiklar skulle därmed kunna bli större. Forskare som däremot främst samarbetar med andra forskare inom den egna disciplinen, har dels heltäckande kunskaper om arbetet som sådant, dels och kanske framförallt, god kännedom om vilka kollegerna är.

I anslutning till dessa tankar om personkännedom vill jag tillägga att det finns andra accepterade sätt för forskare att bedöma ett materials trovärdighet, utöver traditionell peer review. Halliday (2001) nämner läsarens kännedom om den aktuella forskaren, tidskriftens namn och det rykte forskaren, respektive den institution han verkar vid, har. Även vem som står bakom eventuella anslag kan säga något om ett vetenskapligt resultatets trovärdighet, menar Halliday. Via en hänvisning till Kling och McKim (1999) anser Halliday att ”tillförlitlighet bestäms av en kombination av institutionaliserade vanor och läsarens personliga kunskap. I frånvaro av den förra måste forskare förlita sig på den senare” (Halliday 2001; *min översättning*). Kling och McKim menar t.o.m. att en ”icke-granskad ’posting’ på en webbplats, av en välrespekterad forskare med hög status, kan anses mer tillförlitlig än en granskad tidskriftsartikel av någon mindre känd person inom forskarkollektivet” (1999, s. 899; *min översättning*).

Rodman (2003) nämner att det även finns andra ”objektivt mätbara kvalitetskriterier” för att bedöma vetenskaplig tyngd. Hit hör bibliometriska metoder, impact factors (IF) och citeringsanalys. En relativt nytt mått för att uppskatta webbsidors impact factor är det som kallas Web Impact Factor (WIF). Med hjälp av WIF kan man mäta i vilken utsträckning det finns länkar till en webbsida. WIF är ett mått analogt med citeringsanalys inom tryckta medier och liknar även PageRank, Googles metod för att mäta hur populära olika webbsidor är (Smith 2002, s. 134). Troligtvis kommer vi med de elektroniska mediernas utveckling att få se nya metoder för att mäta vetenskaplig tyngd.

5.3.3. Upphovsrätt

En annan förmodad stötesten vid egenarkivering är frågan om upphovsrätt. När forskare publicerar monografier behåller de upphovsrätten och belönas därmed, åtminstone i vissa fall, inte bara med inflytande utan även ekonomiskt. Vid publicering i traditionella vetenskapliga tidskrifter överlåter författaren upphovsrätten till förlagen, i utbyte mot det inflytande traditionell publicering oftast innebär. Överlåtelse av upphovsrätten innebär ofta vissa begränsningar för författarna. Vissa förlag förbjuder uttryckligen fri spridning/arkivering av i förlagets tidskrifter publicerade artiklar (Rabow 2001, s. 39). Likaså vägrar vissa förlag, såväl kommersiella som icke-kommersiella, att acceptera artiklar som tidigare arkiverats i ”öppna arkiv” eller liknande (se t.ex. Kling & McKim 1999; 2000). Dessa omständigheter kan naturligtvis bidra till att forskare inte riktigt vet hur de ska förhålla sig till egenarkivering av pre-prints och granskade artiklar.

Ett försök att bringa klarhet i olika frågor kring egenarkivering och upphovsrätt är det brittiska RoMEO-projektet (Gadd, Oppenheim & Proberts 2003). De tio följande styckena är ett referat av Gadd, Oppenheim och Proberts (2003). RoMEO står för Rights

Metadata for Open-archiving och det nu avslutade projektet ”hade till uppgift att undersöka alla frågor relaterade till den intellektuella upphovsrätten i förhållande till brittiska akademikers öppen-arkivering av forskningsresultat” (s. 1; *min översättning*). Projektets huvudsakliga mål är att producera några metadataelement, med vars hjälp författare kan beskriva ”rättigheterna för deras 'give-away literature' och på så sätt erbjuda den visst skydd i en omgivning av 'open access'” (s. 1; *min översättning*). Den studie jag här hänvisar till här publicerades i mars 2003 och är den första av de sammanlagt sex studier som projektet är tänkt att mynna ut i. I denna studie undersöks hur universitet, författare och förlag ser på upphovsrätten till forskningsresultat.

RoMEO Studies 1 har genomförts med hjälp av en litteraturstudie, en användarundersökning och en analys av 80 tidskriftsförläggares regler för överlåtelse av upphovsrätt (copyright transfer agreements/CTA). Författarna till studien slår fast att underlagen, till både användarundersökningen och analysen av förlagens CTA, är alltför begränsade för att kunna generera några ”avgörande bevis”, men att resultaten ger ”en stark indikation på rådande åsiktsläge” (s. 26; *min översättning*). Av de tre metoder som använts i *RoMEO Studies 1* är användarundersökningen av störst intresse för min uppsats.

Målgruppen för användarundersökningen var akademiska författare och målet med undersökningen var

- 1) To understand academics current practices with regards to the creation of, and the copyright in, research papers;
- 2) To understand academics current practices and views regarding author self-archiving;
- 3) To understand how academics wished to protect their own freely-available research papers on the web with a view to informing the development of some rights metadata elements for this purpose;
- 4) To understand how academics used other people's freely available research papers on the web. (s. 2)

Ett frågeformulär som täckte in dessa fyra områden förmedlades via ett antal diskussionslistor och *arXiv.org e-Print archive*. Frågeformuläret var indelat i tre sektioner: ”A. About You”, ”B. Your research papers” och ”C. How you use other peoples' research papers”. *RoMEO Studies 1* redogör för sektion A och delar av sektion B (s. 2).

Sammanlagt 540 personer svarade på formuläret, av dem kom ungefär hälften från Storbritannien och USA, 176 respektive 92 respondenter. Eftersom formuläret skickades ut on-line är det omöjligt att fastställa någon svarsfrekvens. 50 % av respondenterna kom från naturvetenskapliga discipliner, 38 % från humaniora och samhällsvetenskap och 12 % från ingenjörsvetenskaperna. Sett utifrån inom vilken disciplin respondenterna hörde hemma förekom fysik, biblioteks- och informationsvetenskap samt discipliner inriktade på affärsverksamhet mest frekvent (59, 59 respektive 54 respondenter). Enligt min bedömning speglar resultatet till viss del valet av diskussionslistor där formuläret annonserades.

Som tidigare nämnts överlåter författaren vid publicering av vetenskapligt material i traditionella tidskrifter vanligtvis upphovsrätten till förlaget. I *RoMEO Studies 1* säger det stora flertalet av respondenterna, 485 av 540, att de överlåter upphovsrätten i utbyte mot publicering. Mer än hälften av dem gör det dock motvilligt. *RoMEO Studies 1* ger inget riktigt konkret svar på varför så många av respondenterna känner motvilja mot att lämna över upphovsrätten, men antyder att det har att göra med det förbud mot egenarkivering som överlåtelsen i vissa fall innebär. Författarna till *RoMEO Studies 1* hänvisar till två andra studier, (Hunt 2002) och Swan (2002), vars resultat också tyder på att en bibehållen rätt till egenarkivering är betydelsefull för åtminstone en del författare. Gadd, Oppenheim och Probetz menar att i de fall en författare måste lämna över upphovsrätten, bör han eller hon ta reda på om egenarkivering är tillåten. Om egenarkivering inte är tillåten, bör författaren förhandla sig till en sådan rätt. Denna process är enligt Gadd, Oppenheim och Probetz komplex och avhängig de olika förlagens CTA och författarnas behov.

I Storbritannien har frågan kring upphovsrätt ytterligare en dimension Där har man under senare år börjat diskutera *vem* som egentligen äger upphovsrätten till de forskningsresultat som produceras vid ett universitet. I *The UK Copyright Designs and Patents Acts of 1998* sägs följande:

11. - (2) Where a literary, dramatic, musical or artistic work is made by an employee in the course of his employment, *his employer is the first owner of any copyright* in the work subject to any agreement to the contrary [min kursivering]. (se Gadd et al.2003, s. 7)

Denna lag innebär, enligt vissa, att det är universiteten som äger de resultat de anställda producerar, men att universiteten av tradition har lämnat ifrån sig upphovsrätten till författarna, som i sin tur överlåter upphovsrätten till förlagen i utbyte mot publicering. Att det egentligen är universiteten som äger forskarnas resultat är dock ingen okontroversiell fråga. På frågan "At your institution who owns the copyright in your research papers"? svarade 61 % av respondenterna att det är akademikerna som innehar upphovsrätten, 32 % svarar att de inte vet och bara 7 % att det är institutionen. Vid universitetet i Cambridge har ett försök att införa en ny policy angående upphovsrätten, till universitetets fördel, lett till uppror bland de anställda (s. 13f).

Frånsett denna intrikata fråga visar *RoMEO Studies 1* att i den nuvarande modellen (där upphovsrätten vanligtvis överlämnas av författarna till förlagen) är 50 % av de författare och institutioner som vill egenarkivera förhindrade att göra så. Gadd, Oppenheim och Probetz visar på några alternativa modeller. Den bästa lösningen, enligt Gadd, Oppenheim och Probetz, vore en modell där universitetet behåller upphovsrätten och sedan via en slags licenser överlåter publiceringsrätten till författarna, som sedan i sin tur överlåter publiceringsrätten till olika förlag. Enligt Gadd, Oppenheim och Probetz kan detta vara en modell som inte alla författare accepterar. De lanserar därför en alternativ modell i vilken universitet och författare tillsammans innehar upphovsrätten och sedan på licens överlåter publiceringsrätten till olika förlag.

Gadd, Oppenheim och Probetz visar, anser jag, att oklarheter kring upphovsrätten i stor utsträckning påverkar forskares inställning till "öppna arkiv" och andra alternativ. *RoMEO Studies 1* säger tyvärr ingenting om eventuella olikheter mellan discipliner, i

frågan om och hur upphovsrättsliga frågor påverkar viljan till egenarkivering. Eventuella skillnader mellan discipliner är troligtvis inte den fråga som RoMEO-projektet i första hand är intresserad av.

Hunt (2002) har i en *Archive User Survey* undersökt forskares inställning till arkivering i ”öppna arkiv”. De fem följande styckena bygger på Hunt (2002). Hunt är verksam vid universitetet i Southampton, the Electronic and Computer Science Departement. Bakom undersökningen står bl.a. Steven Harnad, en av de främsta förespråkarna för alternativa modeller inom vetenskapliga kommunikationen. Hunt intresserar sig för fyra användargrupper. Dessa är användare, respektive ickeanvändare, av ett arkiv för fysiker, *arXiv.org e-Print archive* (arXiv.org e-Print archiv 2003), skapat 1991 och användare, respektive ickeanvändare, av *Cogprints* (Cogprints 2003), ett arkiv för papers inom psykologi, neurologi, lingvistik, datavetenskap, filosofi och biologi, skapat 1997. En uppmaning att delta i undersökningen publicerades på ett antal webbplatser, som man förmodade besöktes av akademiker inom de utvalda disciplinerna, bl.a. *arXiv.org e-Print archive* och *Cogprints* webbplatser. Via en länk kunde intresserade ta sig till en webbplats där de, beroende på vilken av de fyra grupperna de ansåg sig tillhöra, kunde välja ett av fyra olika formulär. Varje formulär innehöll 72 frågor. Totalt svarade 664 personer på enkäten, övervägande delen tillhörde användare av *arXiv.org e-Print archive* (389 respondenter) och ickeanvändare av *Cogprints* (166 respondenter). Ett mycket stort material att bearbeta m.a.o. och Hunt medger att det var problematiskt att bestämma vad som var intressant att diskutera (s. 26). Målet med undersökningen var att ta reda på vilka användarna av *arXiv.org e-Print* och *Cogprints* är och hur de använder arkiven. Likaså ville Hunt ta reda på mer om dem som inte använder arkiven och vad som skulle kunna få dem att använda arkiven i större utsträckning.

Några av de resultat Hunt presenterar är följande:

De som arkiverar i *arXiv.org e-Print* har arkiverat sedan en längre tid tillbaka och i större omfattning. De låter sig inte avskräckas av frågor med anknytning till upphovsrätt och plagiering. De arkiverar i synnerhet pre-prints, men även arkivering av post-prints förekommer. De tycks överlag ha ”glömt bort” fördelarna med egenarkivering, den ses som alldeles självklar. Trots att de skulle föredra helt fri publicering har ingen av dem slutat att publicera sig i traditionella tidskrifter.

Användarna av *Cogprints* arkiverar huvudsakligen granskade post-prints. De är medvetna om vikten av synlighet både genom arkivering och genom peer-review. De kan dock inte se fördelarna och vikten av egenarkivering av pre-prints och känner, enligt Hunt, onödig oro över frågor som har att göra med upphovsrätt och plagiering.

Bland dem som inte arkiverade vare sig pre- eller post-prints var de främsta orsakerna en osäkerhet om hur arkiven ska användas och en osäkerhet kring upphovsrättsliga frågor.

Med tanke på det digra material Hunt har haft till sitt förfogande finner jag hennes slutsatser förhållandevis magra. Hon lyckas t.ex. inte ge något riktigt uttömmande svar på *varför* fysiker i större utsträckning använder arkiv än andra forskare. Hon spekulerar i inledningen om möjliga orsaker och nämner följande: Fysiker delade med sig av pre-prints långt innan *arXiv.org e-Print* grundades; fysiker arbetar inom en TeX-kultur, ett format som underlättar utbyte av information; fysiker är mer beroende av andras

resultat; fysiker är mer seriösa med sin forskning och i större behov av snabb spridning av resultat; skillnaden mellan pre- och post-prints är minimal. Vad Hunt menar med att fysiker är mer seriösa framgår inte direkt, men kan liksom Kling och McKims uttalande kanske vara ett utslag av någon slags omedveten värdering. Spekulationerna i uppsatsens inledning berör inte Hunt i någon större utsträckning i sina slutsatser. Enligt Hunt bekräftar dock hennes rapport ”att det är en myt att arkivering enbart passar fysiker och icke-granskade pre-prints och att det på grund av upphovsrättsliga spörsmål inte är möjligt att arkivera. Dessa missuppfattningar kommer att försvinna så fort författare inser fördelarna med allmän on-linepublicering som ett komplement till granskade tidskrifter” (2002; *min översättning*). Här anser jag att Hunt i allt för stor utsträckning begränsar problemet till en bristande insikt hos författare/forskare. Med tanke på att det är Steven Harnad, en av de mer framträdande förespråkarna för ”öppna arkiv”, som ligger bakom undersökningen, är kanske Hunts slutsatser föga överraskande.

Enligt min bedömning tycks oklarheter kring upphovsrätt i varierande grad påverka forskares inställning till ”öppna arkiv” och andra alternativa modeller. Det är svårt att utifrån min empiri avgöra huruvida denna faktor påverkar humanister och samhällsvetares inställning i större utsträckning än naturvetare. Hunts undersökning antyder en skillnad mellan användare av *arXiv.org e-Print* och användare av *Cogprints*. Men hennes resultat kan knappast tas som en intäkt för att humanister och samhällsvetare påverkas mer av oklarheter kring upphovsrätt än naturvetare. *Cogprints* kan inte ses som ett rent humanistiskt eller samhällsvetenskapligt arkiv, utan är snarare ett arkiv för discipliner inom naturvetenskap, humaniora och samhällsvetenskap med inriktning på kognitiv forskning. Hunts studie visar dock att vissa forskargrupper oroar sig mer för upphovsrättsliga frågor än andra. Gadd, Oppenheim och Probetz berör inte skillnader mellan discipliner överhuvudtaget, men även de visar att oklarheter kring upphovsrätten påverkar forskares inställning till egenarkivering. Eftersom många forskare uttrycker en oro kring upphovsrätt och risk för plagiering och olika förlags policy dessutom skiljer sig åt, anser jag att frågor kring upphovsrätt bör få stort utrymme i debatten kring alternativa publiceringsmodeller.

5.3.4. Val av publiceringsform

Den dokumenttyp som oftast nämns i debatten kring alternativa publiceringsformer är tidskriftsartiklar (pre- och post-prints). Ylva Lindholm-Romantschuk är en forskare som har undersökt vilken dokumenttyp forskare inom humaniora och samhällsvetenskap föredrar att använda. Lindholm-Romantschuk doktorerade 1994 vid University of California, Berkeley. Hennes avhandling handlar om informationsflöden inom samhällsvetenskap och humaniora. Enligt Lindholm-Romantschuk (1998, s. vii) tycks det råda brist på studier av vetenskaplig kommunikation inom samhällsvetenskap och humaniora. Genom sin egen forskning försöker hon råda bot på denna obalans. Även monografisk litteratur har enligt Lindholm-Romantschuk ”tenderat att vara relativt negligerad i studier av vetenskaplig kommunikation” (1996, s. 390; *min översättning*). Lindholm-Romantschuk tror att detta kan hänga samman med att forskare företrädesvis studerar kommunikation inom naturvetenskap, där tidskriftsartiklar anses dominera som publiceringsformen.

Som en del i ett större forskningsprogram publicerade Lindholm-Romantschuk för sju år sedan en studie av monografins roll inom vetenskaplig kommunikation (1996, s. 389ff).

Lindholm-Romantschuk undersöker hur idéer sprids inom tre utvalda discipliner inom humaniora och samhällsvetenskap - filosofi, sociologi och ekonomi. Anledningen till att hon valde just dessa tre discipliner, är att de visar på det spektrum av både hårda (ekonomi) och mjuka (filosofi) discipliner som finns inom humaniora och samhällsvetenskap, med sociologi någonstans mitt emellan. Fokus för studien var monografins roll i den vetenskapliga kommunikationen inom nämnda discipliner, i förhållande till den roll vetenskapliga tidskrifter spelar. Lindholm-Romantschuk arbetar utifrån hypotesen att inom vissa discipliner spelar monografier en viktigare roll i spridningen av idéer än tidskriftsartiklar. Som metod använder sig Lindholm-Romantschuk av citeringsanalys. Enligt Lindholm-Romantschuk kan citeringsanalys, trots "välkända metodologiska brister", mycket väl användas vid utvärdering av olika forskningsaktiviteter, i synnerhet om metoden används tillsammans med kvalitativt inriktade metoder (1996, s. 390; *min översättning*).

I sin doktorsavhandling använde sig Lindholm-Romantschuk av ett urval av monografier inom humaniora och samhällsvetenskap, sammanlagt 1600 titlar. Från detta dataset valde hon för denna studie ut de 116 monografier som hörde hemma inom ovan nämnda tre discipliner. En analys av antal citeringar till dessa 116 monografier genomfördes. Likaså genomfördes en analys av alla citeringar till artiklar, skrivna av författarna till de utvalda monografierna under en femårsperiod. Avslutningsvis jämförde Lindholm-Romantschuk data från de två citeringsanalyserna. Resultatet visar att monografier genererar ett större antal citeringar inom alla de tre disciplinerna. Proportionen mellan citeringar till monografier och citeringar till tidskriftsartiklar var för filosofi 7.7, för sociologi 2.6 och för ekonomi 2.4. Hypotesen, att monografier har större betydelse än tidskriftsartiklar inom de tre disciplinerna, kunde därmed anses verifierad. Studien genererade ytterligare några resultat relaterade till monografier, bl.a. att det inom alla tre discipliner finns en kärna av författare, som genererar fler citeringar än övriga och att dessa författare citeras under en lång period, som längst upp till 18-19 år inom sociologi (Lindholm-Romantschuk 1996, s. 389ff).

Delar av denna studie ger, anser jag, kunskap som till viss del kan förklara vissa forskares avvaktande hållning gentemot "öppna arkiv". I den diskussion som pågår kring en omstrukturering av den vetenskapliga kommunikationen, läggs tyngdpunkten på alternativ publicering av tidskriftsartiklar. För discipliner där monografien dominerar som publiceringsform kanske inte egenarkivering framstår som ett lika självklart alternativ, som för de discipliner där artiklar, av tradition, dominerar som publiceringsform. Det finns, som tidigare nämnts, de som menar att "open-access-modellen" passar alla typer av material, men enligt min bedömning är dessa i minoritet. Kan hända är detta enbart en fråga om tid, d.v.s. att i takt med att olika alternativ utvecklas kommer även andra dokumenttyper att tas med i diskussionen.

Det är också intressant med den långa livslängd som åtminstone kärnan av monografier i Lindholm-Romantschuks studie uppvisar. En av fördelarna med "öppna arkiv" sägs vara den *snabbhet* med vilken resultat kan göras tillgängliga - troligtvis en mindre viktig egenskap inom discipliner där resultat står sig i upp mot 20 år. Monografien är också en publiceringsform som väl motsvarar de behov som finns inom många humanistiska och samhällsvetenskapliga discipliner, något som The Knight Higher Education Collaborative (2002, s. 206) påpekar. Författarna menar att humanistiska och samhällsvetenskapliga forskningsresultat ofta har en omfattning som gör att det passar bra att redovisa dem i bokform.

I boken *Scholarly book reviewing in the social sciences and humanities* fortsätter Lindholm-Romantschuk (1998) att undersöka vetenskaplig kommunikation inom humaniora och samhällsvetenskap. Denna gång genomför hon undersökningen med hjälp av recensioner av vetenskapliga monografier. Recensioner har samma funktion för monografier som peer review har för tidskriftsartiklar, d.v.s. en slags kontroll- och legitimitetsfunktion. Utifrån 13 924 recensioner av 1732 monografier undersöker Lindholm-Romantschuk, genom att studera var och hur ofta en bok från disciplin A recenseras inom disciplin B, hur information flödar inom och mellan humanistiska och samhällsvetenskapliga discipliner. En av flera slutsatser som Lindholm-Romantschuk drar är att några discipliner visserligen är mycket isolerade (t.ex. musik och dans), men att det stora flertalet discipliner inom humaniora och samhällsvetenskap är mer eller mindre öppna för informationsutbyte över gränserna. Denna typ av öppenhet är, enligt Lindholm-Romantschuk, mindre vanlig inom naturvetenskapliga discipliner. Hon fastslår dock att det flödar mer information mellan discipliner inom samma grupp än mellan de olika grupperna. När det gäller utbyte av information mellan grupperna flödar det mer information från samhällsvetenskap till humaniora än tvärtom. Lindholm-Romantschuk menar att hennes studie visar att gränserna mellan discipliner "kanske är mer genomträngliga än vad man tidigare trott" (1998, s. 131; *min översättning*). Denna studie bidrar inte direkt till något svar på frågan varför vissa forskare förhåller sig avvaktande till "öppna arkiv". Däremot tyder resultatet på, anser jag, att det finns en relativt stor öppenhet mellan discipliner inom humaniora och samhällsvetenskaperna. Dessa forskningsområden skulle därmed kunna dra nytta av "öppna arkiv", som bland annat sägs underlätta just tvärvetenskaplig kommunikation.

5.3.5. Tekniska faktorer

Rubriken ovan kan inbegripa flera olika aspekter. Övergången till elektroniska medier inom olika forskargrupper kan t.ex. påverkas både av tillgången till ny teknik och av forskares inställning till att börja använda ny teknik. Costa och Medaows (2000), verksamma vid Loughborough University, Storbritannien, har utifrån båda dessa aspekter undersökt hur datorer och Internet har påverkat kommunikationen bland ett antal samhällsvetenskapligt inriktade forskare. I inledningen slår de fast att tidigare studier visat att informationsbehoven ser olika ut inom naturvetenskap och samhällsvetenskap. Ett exempel är att tidskrifter och monografier är viktiga informationskällor inom båda forskningsområdena, men att samhällsvetare i större utsträckning använder sig av monografier. De menar också att det finns tydliga skillnader även mellan olika samhällsvetenskapliga discipliner. Som exempel nämner de att monografier är mindre viktiga för ekonomer än för sociologer. När det gäller användningen av datorer och nätverk visar ett flertal studier på en obalans mellan naturvetenskap och samhällsvetenskap och mellan samhällsvetenskap och humaniora. Men, påpekar Costa och Meadows, forskare inom naturvetenskap "fick tillgång till datorer och nätverk långt tidigare än samhällsvetenskapligt inriktade forskare" (2000, s. 256; *min översättning*). De fastslår också att skillnaderna mellan grupperna är på väg att försvinna (2000, s. 256). Costa och Meadows nämner inte "öppna arkiv" i sin artikel, men drar slutsatser som, enligt min mening, kan öka förståelsen för olika attityder till egenarkivering.

I undersökningen ingick en grupp brasilianska forskare, samt en mindre grupp brittiska forskare, inom två discipliner, sociologi och ekonomi. Disciplinen sociologi valde Costa och Meadows därför att den tillhör "den centrala kärnan av samhällsvetenskap" och

ekonomi för att den tillhör de mer hårda disciplinerna inom samhällsvetenskap (2000, s. 255; *min översättning*). De valde att studera forskare från Brasilien därför att tidigare studier nästan uteslutande har riktats mot datorbaserad kommunikation i industrialiserade länder.

Undersökningen genomfördes i två steg. Det första steget innebar att 760 utvalda brasilianska akademiker, ungefär hälften från varje disciplin, fick svara på en enkätundersökning. Med tanke på att strejk utbröt vid en del av de utvalda universiteten direkt efter att enkäten skickats ut, var svarsfrekvensen förhållandevis god, 64.1 %. Enkäten syftade till att ge en "semikvantitativ" översikt. I steg två, den kvalitativa delen av undersökningen, djupintervjuades 16 ekonomer och 20 sociologer. Steg två föregicks av en pilotstudie där elva brittiska akademiker intervjuades.

Resultaten visar, menar Costa och Meadows, att "även om traditionella informationskällor och tjänster fortsätter att dominera är deras elektroniska motsvarigheter på väg att hinna i fatt" (2000, s. 257; *min översättning*). Förändringen mot en högre användning av elektroniska medier hade dock nått längre bland ekonomerna. Både enkäten och intervjuerna visar dessutom tydligt att "ekonomer har en mer positiv attityd än sociologer gentemot nuvarande och framtida användning av IT" (Costa & Meadows 2000, s. 260; *min översättning*). Denna skillnad kan, enligt Costa och Meadows, delvis förklaras med att ekonomerna vid undersökningstillfället hade betydligt större tillgång till nätverksanslutna datorer. Forskare inom båda disciplinerna var dock överens om att IT kommer att leda till förändringar av den vetenskapliga kommunikationen. Bl.a. håller en stor del av dem med om att den nya tekniken kommer att leda till ökad informell kommunikation, ökat antalet publikationer, förbättrad kvalitet samt att IT kommer att öka och underlätta samarbete både inom och mellan discipliner. Något mer tveksamma var båda grupperna till om tekniken kommer att leda till ökad publicering i elektroniska tidskrifter. De kände ingen större press från forskarkollektivet att själva publicera sig i elektronisk form, men i synnerhet ekonomerna kände en ökad press från omgivningen att läsa elektroniska publikationer.

Undersökningen visar att det även bland forskare inom samhällsvetenskap håller på att ske stora förändringar inom vetenskaplig kommunikation och då i synnerhet när det gäller informell kommunikation. Vissa discipliner inom samhällsvetenskap har, jämfört med andra, varit snabbare på att börja använda den nya tekniken. Undersökningen visar dock en överlag positiv attityd till elektroniska medier, både bland ekonomer och bland sociologer. Ett antal av respondenterna menade också att en ökad användning av elektronisk kommunikation leder till en "demokratisering av det samhällsvetenskapliga forskarsamhället" (Costa & Meadows 2000, s. 261; *min översättning*), med mer jämlik tillgång till information för forskare i utvecklingsländerna och ett ökat utbyte länder emellan. Jag anser att Costa och Meadows undersökning visar på en trolig delförklaring till att egenarkivering hitintills har dominerat inom naturvetenskapliga discipliner. Forskare inom naturvetenskap fick tillgång till den teknik som egenarkivering bygger på, både tidigare och i större omfattning jämfört med forskare inom humaniora och samhällsvetenskap. Bland naturvetare har datorer dessutom i större utsträckning varit, inte bara ett kommunikationsmedium, utan ett arbetsredskap i sig. Idag ser vi en tydlig förändring mot att datorer blir ett arbetsredskap även för humanister och samhällsvetare. Jag tror att denna förändring av tillgång och inställning till tekniken i sig, troligtvis kan öppna upp för en mer positiv attityd gentemot alternativa modeller även inom humaniora och samhällsvetenskap.

5.3.6. Ekonomiska faktorer

I essän *Op.cit.: publishing in the humanities and social sciences* (The Knight Higher Education Collaborative 2002) diskuteras digitala medier och alternativ publicering utifrån bl.a. ekonomiska aspekter. Essän är resultatet av en konferens kring vetenskaplig publicering inom humaniora och samhällsvetenskap. Konferensen kom till som ett svar på en annan konferens där krisen inom vetenskaplig publicering också hade diskuterats, men då med fokus på naturvetenskap, medicin och teknik. Avsikten med den nya konferensen var att diskutera vetenskaplig kommunikation och publicering inom samhällsvetenskap och humaniora, i ljuset av kommunikation och publicering inom naturvetenskap, medicin och teknik. Essän utgår visserligen från amerikanska förhållanden, men ger en bild av humaniora och samhällsvetenskap som nog, åtminstone delvis, kan anses gälla även för vår del av världen.

The Knight Higher Education Collaborative (2002, s. 209; *min översättning*) menar att bristen på pengar inom humaniora och samhällsvetenskap bidrar till att dessa discipliner inte lika snabbt utvecklat "alternativa kanaler för publicering och spridning" av vetenskapligt material, jämfört med naturvetenskapliga och tillämpade vetenskaper där omfattande bidrag utifrån underlättat det samma. Borgman (2000, s. 89) nämner att Ginsparg, när han skulle bygga upp sin e-printserver för fysiker, tog emot ett bidrag på 1 miljon dollar för mjukvaruutveckling. Dessa summor gör att Borgman, även om hon menar att "Ginspargs succé måste prisas", känner tvekan inför modellens "överförbarhet till andra områden" (Borgman 2000b, s. 89; *min översättning*). Det är, enligt min bedömning, svårt att avgöra i vilken utsträckning tillgång/brist på pengar kommer att styra utvecklingen av alternativa publiceringsmodeller. Numera finns det gratis mjukvaruprogram att tillgå för dem som vill bygga upp ett arkiv. Även kostnader för t.ex. långsiktigt underhåll och teknikutveckling måste dock tas med i beräkningen.

Essän ger en bild av ett tudelat samhälle där forskning inom naturvetenskap, teknik och medicin premieras, på bekostnad av forskning inom samhällsvetenskap och humaniora. Denna tudelning kommer huvudsakligen till uttryck genom samhällets fördelning av forskningsanslag. Den ojämlika fördelningen av pengar är inte bara en rent ekonomisk fråga, menar författarna, utan ger även uttryck för samhällets nedvärderande syn på humaniora och samhällsvetenskap. The Knight Higher Education Collaborative menar, som tidigare nämnts, att snedfördelningen av anslag kan ses som en möjlig orsak till den långsammare utvecklingen av alternativa modeller inom humaniora och samhällsvetenskap. En annan orsak kan vara ett "kulturellt betingat motstånd mot det digitala mediet" som många, enligt The Knight Higher Education Collaborative (s. 211; *min översättning*), upplever finns inom humaniora och samhällsvetenskap och de ger en bild av två, något isolerade och defensiva forskningsområden som lite uppgivet betraktar världen. The Knight Higher Education Collaborative understryker att det krävs ett ökat finansiellt stöd och ett ökat socialt erkännande av humaniora och samhällsvetenskap. Men The Knight Higher Education Collaborative menar även att förändringen måste komma inifrån disciplinerna själva. De nämner några olika digitala projekt, som "kan stå modell för ett bredare kulturellt skifte och bidra till att ge humaniora och samhällsvetenskap en ökad närvaro och betydelse generellt i samhället" (s. 212; *min översättning*). Även inom humanistisk och samhällsvetenskaplig kommunikation är ledorden tillgänglighet, effektivitet och tillit. The Knight Higher Education Collaborative menar att, likväl som inom naturvetenskap, kan "en genomtänkt och samlad ansträngning för att utveckla det digitala mediet som

publikationsform” leda till att dessa mål uppfylls inom humaniora och samhällsvetenskap (s. 212; *min översättning*).

The Knight Higher Education Collaborative syn på humaniora och samhällsvetenskap, som två något isolerade och defensiva forskningsområden, motsägs till viss del av Lindholm-Romantschuks undersökning. Enligt henne kan det finnas en större öppenhet inom dessa områden än vad man tidigare har trott. Costa och Meadows undersökning visar också att tydliga förändringar är på gång inom samhällsvetenskap när det gäller forskares inställning till digitala medier. Enligt mig bidrar The Knights Higher Education Collaborative, trots den något negativa tonen, med viktiga synpunkter. De visar att även humaniora och samhällsvetenskap kan och bör, närma sig de digitala medierna. Förändringen måste dock komma både utifrån och inifrån disciplinerna själva. Det handlar inte om att anamma den nya tekniken för teknikens egen skull, utan för att de digitala medierna kan bidra till att synliggöra humanistisk och samhällsvetenskaplig forskning.

5.4. Sammanfattning avsnitt 5

Det pågår en intensiv debatt om hur framtidens vetenskapliga kommunikation kommer att se ut. Vissa menar att en stor del av den vetenskapliga kommunikationen, såväl informell som formell, kommer att ske i elektronisk form, t.ex. Harnad (2001) och Odlyzko (2002). Andra förhåller sig mer skeptiska och menar att olika discipliner kommer att närma sig de digitala medierna på olika sätt och i olika utsträckning (Borgman 2000b; Hurd 2000; Kling & McKim 2000; Meadows 1998). ”Öppna arkiv” har inte heller blivit den omedelbara succé som många hoppades på. Enligt vissa debattörer har forskare inom humaniora och samhällsvetenskap, jämfört med forskare inom naturvetenskap, förhållit sig tämligen avvaktande till alternativa modeller för vetenskaplig publicering (Hunt 2002; Tomaiuolo & Packer 2000; The Knight Higher Education Collaborative 2002). Även bland vissa forskare inom naturvetenskap tycks det finnas en avvaktande attityd (Kling & McKim 1999). I litteraturen framkommer flera tänkbara orsaker till att det förhåller sig så.

Den kanske viktigaste orsaken är att idén med ”öppna arkiv” bygger på en kommunikationsmodell som har sitt ursprung inom ett naturvetenskapligt forskarkollektiv. Den kommunikationsmodellen, uppbyggd kring olika e-printarkiv, skapades utifrån olika behov inom just detta forskarkollektiv. Ett av dessa behov var att snabbt kunna dela med sig av vetenskapliga resultat i form av elektroniska pre-prints. Det har visat sig svårt att överföra denna modell, som så tydligt bygger på ett forskarkollektivs särskilda behov, till forskarkollektiv med helt andra behov (se t.ex. Kling & McKim 2000). Forskares sätt att kommunicera skiljer sig åt, det har åtskilliga studier visat. Jag har i min uppsats presenterat några få av dem (Meadows 1998; Costa & Meadows 1999; Lindholm-Romantschuk 1996; 1998). En av dessa studier (Lindholm-Romantschuk 1996) visar tydligt att inom humaniora och samhällsvetenskap dominerar monografin som kommunikationsform. Detta kan, enligt min bedömning, vara en orsak till just humanistiska och samhällsvetenskapligt inriktade forskares avvaktande förhållningssätt. I den debatt som förs kring en ny modell för vetenskaplig kommunikation talas det, som jag ser det, betydligt oftare om tidskriftsartiklar, i form av både pre-prints och post-prints, än om monografier. Vissa debattörer nämner att den nya kommunikationsmodellen passar för alla typer av vetenskapligt material (Johnson 2002). Dessa debattörer är dock enligt min bedömning, i minoritet. Jag menar att den

nya modellen för vetenskaplig kommunikation, med sitt ursprung inom naturvetenskaplig kommunikation och med sin betoning på tidskriftartiklar, troligtvis framstår som mer främmande för de forskare som huvudsakligen förmedlar sina idéer via monografier.

En annan faktor som tycks påverka många forskare negativt, och då inte bara de inom humaniora och samhällsvetenskap, är oklarheter kring upphovsrätten. RoMEO-projektet (Gadd, Oppenheim & Proberts 2003) visar på ett tydligt sätt hur komplicerad denna fråga är. Många forskare känner rädsla för att arkivering i ”öppna arkiv” kan påverka relationen till en framtida förläggare negativt. Likaså förbjuder vissa förläggare uttryckligen arkivering av tidigare publicerade artiklar. Detta är en fråga som, enligt min bedömning, bör få stort utrymme i den framtida diskussionen kring alternativa arkiverings- och publiceringsformer.

Även synen på peer review tycks påverka forskare inställning till. För forskare inom vissa discipliner, t.ex. fysik, där öppenheten av olika anledningar sägs vara stor och där forskare sedan lång tid tillbaka delar med sig av sina resultat i form av pre-prints, tycks frågan om bristen på traditionell peer review inte vara något stort problem (Kling & McKim 2000). Humaniora och samhällsvetenskap ses av vissa som förhållandevis slutna forskningsområden (Kling & McKim 2000; The Knight Higher Education Collaborative 2002), d.v.s. forskare inom dessa discipliner har inte för vana att dela med sig av resultat under arbetets gång. För forskare med ingen eller begränsad vana av pre-prints, tycks traditionell peer review ha större betydelse för tilliten. Denna syn på den ensamme, isolerade forskaren inom humaniora och samhällsvetenskap kan dock ifrågasättas och två andra studier (Sely & Costa 1999; Lindholm-Romantschuk 1998) som jag hänvisar till bidrar med en annan, mer positiv syn på samhällsvetare och humanister. Många discipliner inom humaniora och samhällsvetenskap är till sin natur tvärvetenskapliga (Lindholm Romantschuk 1998), vilket skulle kunna bidra till ett större behov av traditionell kvalitetsgranskning.

The Knight Higher Education Collaborative (2002) pekar på två orsaker till att digitala medier inte används lika flitigt inom humaniora och samhällsvetenskap. För det första prioriterar samhället naturvetenskap, teknik och medicin, på bekostnad av humaniora och samhällsvetenskap. För det andra finns det inom humaniora och samhällsvetenskap ett kulturellt betingat motstånd mot användning av digitala medier. Costa och Meadows (1999) har dock i en undersökning visat att skillnaderna mellan de tre forskningsområdena minskar, både vad gäller tillgången och inställningen till datorer och digitala medier.

6. Diskussion

Jag har i min uppsats beskrivit ett pågående systemskifte inom vetenskaplig kommunikation och närmare undersökt ett av uttrycken för detta systemskifte, *Open Archives Initiative*. Mina två syften har varit att närma mig Open Archives Initiative dels utifrån ett kunskapsorganisatoriskt perspektiv, dels utifrån ett historiskt, socialt och kulturellt perspektiv. För mig är dessa två syften egentligen bara olika sidor av samma mynt. För att förstå det som sker inom vetenskaplig kommunikation av idag behövs det, enligt min mening, kunskap om såväl tekniska som kunskapsorganisatoriska och sociala, historiska och kulturella aspekter.

Jag har haft för avsikt att besvara följande tre frågeställningar:

- Hur fungerar ”öppna arkiv” utifrån ett kunskapsorganisatoriskt perspektiv, d.v.s. vilka metoder används vid lagring, sökning och återvinning av information?
- Vilka eventuella barriärer finns det mot en vidareutveckling av ”öppna arkiv” och andra alternativa publiceringsformer?
- Om forskare inom humaniora och samhällsvetenskap i större utsträckning förhåller sig avvaktande till ”öppna arkiv” än forskare inom naturvetenskap, vilka tänkbara orsaker kan det i så fall finnas till att det förhåller sig så?

Utän tvekan finns det vissa likheter mellan framväxten av den vetenskapliga tidskriften som ny publiceringsform under 1600-talet och dagens situation, där helt nya former för vetenskaplig publicering är på väg att skapas, varav ”öppna arkiv” är en. Naturligtvis är dagens situation oändligt mycket mer komplex, men det är ändå intressant att notera historiska likheter. De grundläggande likheter som går att se, trots att över 300 år har förflutit, är att både ny teknik och nya behov bidrar till framväxten av nya kommunikations- och publikationsformer, något som Meadows (1998) tidigare visat på. Likaså kan vi även idag se tendenser till en ny kunskapssyn, vilket nya tvärvetenskapliga discipliner är ett tecken på. Andersen (2002) menar att ny kommunikationsteknik på sikt kan förändra ett samhälles sociala kunskapsorganisation. Den sociala kunskapsorganisationen tar sig bl.a. uttryck i hur ett samhälle väljer att dela in den vetenskapliga kunskapsproduktionen i olika discipliner. Kanske kan vi förvänta oss att gränserna mellan naturvetenskap, samhällsvetenskap och humaniora blir allt mer flytande i framtiden och då delvis som en följd av att IT ger nya möjligheter till samarbete. Detta skulle på lång sikt kunna få till följd att kunskap måste organiseras på ett nytt sätt.

Den vetenskapliga kommunikationen, d.v.s. de informella och formella samtal som förs mellan olika forskare och mellan forskare och samhället i stort, är en mycket viktig del av all forskning. Det handlar både om att dela med sig av resultat under arbetets gång och att publicera slutresultat. Även om jag anser att det finns stora likheter mellan situationen för trehundra år sedan och den idag, finns det även det som skiljer. Utöver tillväxten inom forskningen under andra halvan av 1900-talet och att datorer, Internet och webben har gett forskare nya kommunikationsmöjligheter, finns det även en finansiell kris bakom framväxten av alternativa publiceringsformer. Bibliotekens ökade kostnader för vetenskapliga tidskrifter de senaste 30 åren, i ett i övrigt kärvt ekonomiskt klimat, har bidragit till att även universitet och högskolor börjat intressera sig för alternativa publiceringsformer.

Ett rad olika alternativ till traditionell vetenskaplig publicering har på senare år sett dagens ljus. Dessa alternativ kan huvudsakligen delas in i ”öppna arkiv” och ”open access journals”. I denna uppsats har jag främst undersökt utvecklingen av ”öppna arkiv”.

”Öppna arkiv” kan vara av tre slag:

- Forskares privata webbplatser

- Ämnesarkiv
- Institutionsarkiv, företrädesvis vid ett universitet eller en högskola

Av dessa tre är det troligtvis ämnesarkiv och institutionsarkiv som på sikt kommer att få störst betydelse. I arkiven ska forskare, på egen hand eller i samarbete med olika informationsspecialister, arkivera/publicera sina resultat, som komplement till eller som ersättning för publicering i traditionella vetenskapliga tidskrifter.

Bakgrunden till utvecklingen av OAI: s återvinningsmodell står att finna inom en särskild forskargrupp, fysikerkollektivet. Fysiker har en lång tradition av att dela med sig av s.k. pre-prints och skapade 1991 det första e-printarkivet. Detta e-printarkiv kan sägas vara föregångaren till dagens ”öppna arkiv”. Några år senare insåg andra forskarkollektiv att även de skulle kunna dra nytta av fysikernas modell och flera nya e-printarkiv etablerades. De som använde arkiven insåg att en möjlighet till samsökning skulle kunna öka arkivens användbarhet och inflytande. I slutet av 1999 bildades OAI med målet att möjliggöra samsökning i olika arkiv. Tekniken i OAI: s modell är förhållandevis enkel och bygger på två koncept, utbyte av metadata enligt formatet Dublin Core och ett protokoll kallat *the OAI Protocol for Metadata Harvesting 2.0*. Med hjälp av protokollets sex frågor samlar informationstjänster in metadata ”en masse” från olika arkiv och lagrar sedan metadatan i databaser, i vilka användarna sedan kan söka information. Tanken är att olika typer av informationstjänster ska byggas upp och dessa ska erbjuda olika tilläggstjänster utifrån den tilltänkta användargruppens behov.

OAI har rönt mycket uppmärksamhet och många hoppas på att modellen ska bidra till ökat samarbete forskare emellan och en ökad tillgänglighet till vetenskapliga resultat. Den teknik som ”öppna arkiv” bygger på sägs fungera tillfredsställande. OAI-PMH arbetades dock fram under mycket kort tid och allteftersom har en del svagheter uppmärksamats. En del av dessa svagheter är förknippade med de mänskliga aspekterna av arbetet. Det har t.ex. uppstått problem vid själva skapandet av metadata, trots att OAI använder sig av Dublin Core, ett metadataformat som är känt för att vara enkelt att använda. Det förs också diskussioner kring *vem* som bör skapa metadata och *hur* den ska skapas. Själva arkiveringsmomentet har också lett till problem i vissa fall. Andra frågor är förknippade med innehållet i arkiven, t.ex. vilka dokumenttyper som ska arkiveras och i vilka format. Även den långsiktiga finansieringen av ”öppna arkiv” är föremål för diskussion.

Den största utmaningen tycks ligga i att få forskare att vilja använda ”öppna arkiv”. Borgman (2000a; 2000b) diskuterar sociala processer inom vetenskaplig kommunikation och intresserar sig bl.a. för hur den sociala kontexten påverkar hur forskare väljer att kommunicera sina resultat. Borgman skriver om vetenskaplig kommunikation utifrån hypotesen att alternativa publiceringsformer är något som vi kommer att få se mer av framöver. Hon menar också att vi behöver mer kunskaper om ”hur, varför, när och för vem människor skapar dokument om vi ska kunna bygga bättre verktyg för dem att göra det” (Borgman 2000b, s. 94). Genom att studera utifrån vilka kriterier forskare väljer publiceringsform kan vi få kunskap som bl.a. de digitala biblioteken kan användas vid skapandet av alternativa publiceringsformer. Som jag ser det har Borgmans tankar hög relevans för det fortsatta arbetet med ”öppna arkiv”. Kunskap om olika forskargrupperns behov, enligt Borgmans modell, är enligt min mening en förutsättning för att ”öppna arkiv” och andra alternativa modeller verkligen

ska bli en ny modell för vetenskaplig kommunikation. Med en fokusering på vad människan ska kunna göra med hjälp av teknik och inte på vem som ska använda den eller varför, riskerar den nya modellen att bli en verklighet för bara vissa forskargrupper och vinsterna därmed begränsade.

Jag har i denna uppsats försökt urskilja ett antal barriärer som kan tänkas försvåra eller försena en övergång till en mer omfattande användning av ”öppna arkiv”. En av mina frågeställningar var också att undersöka huruvida forskare inom humaniora och samhällsvetenskap förhåller sig mer avvaktande än forskare inom naturvetenskap. Även om min empiri inte direkt ger några konkreta bevis för att motståndet är större bland forskare inom humaniora och samhällsvetenskap, visar undersökningen att forskare inom dessa discipliner har varit långsammare med att acceptera ”öppna arkiv” som en ny modell för vetenskaplig kommunikation. Ibland verkar dock skiljelinjen snarare gå mellan å ena sidan *vissa* naturvetenskapliga discipliner och å andra sidan *andra* naturvetenskapliga samt humanistiska och samhällsvetenskapliga discipliner. Brist på traditionell kvalitetsgranskning och oklarheter kring upphovsrätt verkar vara två av de faktorer som i hög grad påverkar forskares inställning till egenarkivering. Det finns dock vissa omständigheter som kan förklara just humanisters och samhällsvetares avvaktande förhållningssätt. En framträdande orsak, enligt min mening, är att ”öppna arkiv” bygger på en modell som utvecklats utifrån speciella behov hos en viss grupp forskare - fysiker. Ett av dessa behov är att mycket snabbt kunna dela med sig av pre-prints. Detta behov av att snabbt kunna förmedla resultat till forskarkolleger är inte tillnärmelsevis lika stort inom humaniora och samhällsvetenskap. En annan möjlig orsak är att ”öppna arkiv” främst ska underlätta arkivering/publicering av tidskriftsartiklar, både pre- och post-prints. Inom humaniora och samhällsvetenskap har monografin av tradition en mycket stark ställning. Det förekommer även publicering i tidskrifter, men som bl.a. Lindholm-Romatschuk (1996) visat, är detta mindre vanligt förekommande, i synnerhet inom humaniora. Ytterligare en förklaring skulle kunna vara att många av disciplinerna inom humaniora och samhällsvetenskap är tvärvetenskapliga. När forskare hämtar idéer från fler områden än det man huvudsakligen är verksam inom, kan behovet av traditionell peer review troligtvis bli större.

Det finns dock fördelar med ”öppna arkiv” som skulle kunna vara skäl nog för forskare inom humaniora och samhällsvetenskap att använda ”öppna arkiv” i större utsträckning. Arkivering/publicering i ”öppna arkiv” skulle på sikt kunna leda till att humaniora och samhällsvetenskapliga resultat når en större publik än vad fallet är idag. The Knight Higher Education Collaborative (2002) menar att det finns en tendens till isolering inom dessa vetenskaper, en isolering som delvis kan bero på forskarna själva. En ökad användning av ”öppna arkiv” och andra alternativ skulle, dels kunna öka det konkreta samarbetet mellan olika discipliner, dels sprida resultaten utanför den egna gruppen. Detta skulle enligt The Knight Higher Education Collaborative vara mycket fruktbart, både för forskarna själva och för samhället i stort. Vissa av de undersökningar jag har hänvisat till visar också att det är förändringar på gång inom humaniora och samhällsvetenskap. Den av vissa påstådda slutenheten inom dessa forskningsområden verkar inte vara så stor som man tidigare trott och en ökad tillgång till datorer öppnar upp för nya kommunikationsmönster även inom dessa discipliner.

En annan intressant, och enligt min mening mycket viktig, aspekt av arbetet med alternativa modeller för vetenskaplig kommunikation, är de digitala bibliotekens roll i detta arbete. Att tala om *digitala* bibliotek blir dock något missvisande. Än så länge är

de digitala biblioteken oftast integrerade i traditionella, fysiska bibliotek, i det som ofta kallas hybridbibliotek. Satsningar på institutionsarkiv blir därmed en uppgift inte bara för de digitala biblioteken, utan för biblioteken som helhet. Nedan kommer jag därför att tala om *bibliotek*, jag avser då högskole- och universitetsbibliotek, varav de digitala biblioteken underförstått är en del.

Biblioteken är en av tre aktörer i den traditionella publiceringskedjan. En av deras främsta roller har traditionellt sett varit att välja ut och i fysisk eller elektronisk form tillhandahålla olika dokumenttyper. I viss mån producerar vissa universitet och högskolor även egen litteratur. Inom den nya modellen för vetenskaplig kommunikation, vars omfattning vi ännu så länge bara kan sia om, kan bibliotekens roll som producenter/förläggare bli betydligt mer omfattande.

Biblioteken har en hel del att vinna på de alternativa modeller för vetenskaplig kommunikation som är på väg att växa fram. Biblioteken har i allra högsta grad drabbats av de prisstegringar på vetenskapliga tidskrifter som kommersiella förlag genomfört sedan början av 1970-talet. Prisstegringar, i kombination med en allt mer begränsad budget, har tvingat biblioteken att välja ut vissa tidskrifter, på bekostnad av andra. En allvarlig följd är att tillgängligheten till olika tidskrifter blir mycket ojämlig. De ökade kostnaderna för vetenskapliga tidskrifter har även fått andra konsekvenser, t.ex. att inköpen av monografier har minskat. Detta blir extra bekymmersamt för forskare och studenter inom humaniora och samhällsvetenskap där monografin dominerar som publikationsform. En ökad satsning på alternativa publiceringsmodeller skulle kunna leda till förbättringar för biblioteken på flera sätt. Dels kan en ny modell leda till att biblioteken kan omfördela sina resurser. I stället för att som idag betala dyrt för prenumerationer på de kommersiella förlagens tidskrifter, kan biblioteken lägga resurser på t.ex. ett institutionsarkiv. Ett institutionsarkiv kan på sikt leda till ökad tillgänglighet, även till litteratur som biblioteken idag inte anser sig ha råd med. Men det är också viktigt, anser jag, att diskutera vad en omfattande satsning på digital publicering kan få för konsekvenser för mer traditionella publiceringsformer, t.ex. tryckta monografier, som av tradition spelar en stor roll inom humaniora och samhällsvetenskap. Vilket utrymme och vilka möjligheter finns det egentligen inom det nya publiceringsparadigmet för digital publicering av monografier? Inom universitetsvärlden börjar e-böcker bli allt vanligare, men marknaden för akademiska e-böcker domineras i hög grad av några få stora leverantörer, t.ex. NetLibrary och Ebrary. Blir dessa e-boksleverantörer nästa stora utmaning för forskare och bibliotek?

Pinfield, Gardner och MacColl (2002) menar att biblioteken har de ekonomiska, tekniska och organisatoriska förutsättningar som behövs vid uppbyggnaden av institutionsarkiv. Vid biblioteken finns också omfattande kunskaper, i t.ex. katalogisering och skapande av metadata, som kan komma att bli mycket användbara. OAI:s modell för samsökning av olika arkiv bygger på metadataformatet Dublin Core. Pinfield, Gardner och MacColl har visat att det kan uppstå problem vid skapandet av metadata. Här skulle, menar jag, bibliotekarier, med sina omfattande kunskaper i katalogisering, klassifikation och indexering, kunna fylla en viktig funktion genom att bidra med sina kunskaper. Även när det gäller själva arkiveringen tror Pinfield, Gardner och MacColl att biblioteken skulle kunna fungera som en slags mellanhand. Vid de universitet och högskolor som väljer att skapa arkiv kan jag se att personalen kan komma att få ytterligare en viktig roll, den att få forskare att vilja använda arkiven. Med tanke på vad olika studier har visat när det gäller inställningen till egenarkivering

kommer olika forskare, och för den delen även studenter, med all sannolikhet att i varierande grad vilja använda arkiven. Sätten att använda dem på kommer säkerligen också att variera.

Crow (2002) har visat på ytterligare en vinst för de universitet som väljer att bygga upp institutionsarkiv. Som det ser ut idag sprids resultaten av den forskning som bedrivs vid ett universitet på ett stort antal tidskrifter. Vissa resultat publiceras i tidskrifter med hög status och bred publik, andra i tidskrifter som når en mycket begränsad publik. Genom ett institutionsarkiv ökar universitetets möjligheter att synliggöra den forskning som bedrivs. Arkivet kan på så vis både öka tillgängligheten för alla slags resultat och bli en kvalitetsindikator som i förlängningen kan leda till en höjning av universitets status.

Ytterligare en aspekt att ta hänsyn till, när det gäller bibliotekens roll i uppbyggnaden av alternativa modeller, är behovet av nya samarbetsformer för olika personalkategorier. Forskare och personal vid biblioteken kommer troligtvis att behöva finna nya former för samarbete, likaså olika grupper vid biblioteken, t.ex. bibliotekarier, informationsvetare, dataspecialister etc. I detta behov av ökat samarbete ligger troligtvis stora utmaningar, men också möjligheter till givande kunskapsutbyte. Både bibliotek och forskare har ett gemensamt intresse av att finna modeller som leder till en omfördelning av bibliotekens resurser, från de kommersiella förlagen till biblioteken och användarna. Likaså har biblioteken vid våra högskolor och universitet en viktig roll att spela och ett ansvar för att olika forskargrupperns behov tillgodoses i lika stor utsträckning.

När jag nu befinner mig i uppsatsens slutskede har jag blivit varse svårigheterna med att inom ramen för en magisteruppsats, med ett begränsat antal sidor till mitt förfogande, ha två syften. Var och ett av mina syften skulle med lätthet kunna fylla en magisteruppsats. I synnerhet kan jag se att avsnittet om OAI skulle ha tjänat på att vara mer omfattande. Det finns utan tvekan många intressanta områden inom OAI: s återvinningsmodell, med anknytning till kunskapsorganisation, som är värda att undersökas. Jag har dock blivit övertygad om att ett kunskapsorganisatoriskt perspektiv med fördel kan kombineras med ett historiskt, socialt och kulturellt perspektiv. Med två syften har jag dock tagit en risk och den risken är att min uppsats saknar djup. Förhoppningsvis har den i stället en bredd, som i någon mån kan bidra till en större förståelse för den framtid den vetenskapliga kommunikationen går till mötes. Uppsatsen har inte heller genererat ny kunskap i lika stor utsträckning som jag i början av uppsatsskrivandet hoppades på. Det kan delvis bero på valet av metod. I en litteraturstudie finns alltid en risk att hamna i ett allt för omfattande refererande. Jag har försökt att kombinera redan genomförda studier och andra texter med nya frågeställningar, men kan se att min empiri inte har varit tillräckligt omfattande för att ge några riktigt bra svar på mina frågeställningar. Jag anser dock se att min uppsats skulle kunna fungera som en förstudie till en empirisk undersökning, kring t.ex. metadataskapande eller olika discipliners inställning till egenarkivering, en studie jag till en början hade tankar på att genomföra men som jag, huvudsakligen av tidsbrist, valde att avstå ifrån.

Vetenskaplig kommunikation genomgår utan tvekan en förändring som vi än så länge bara har sett början av. En starkt bidragande orsak till denna förändring är framväxten av ny informationsteknik och de radikalt annorlunda möjligheter till kommunikation som denna teknik ger möjlighet till. Övergången från tryckta till elektroniska medier har pågått en längre tid och i den förändring som sker kommer Open Archives Initiative och andra aktörer att spela en viktig roll. Huruvida det som sker kan klassas som ett

paradigmskifte inom vetenskaplig kommunikation återstår att se, likaså i vilken omfattning olika discipliner kommer att anamma den nya modellen. Tryckta medier kommer med all sannolikhet att leva kvar jämsides med elektroniska medier under många år framöver. Jag är också övertygad om att olika discipliner kommer att fortsätta närma sig de digitala medierna på olika sätt. Det är mycket möjligt att forskare inom humaniora och samhällsvetenskap kommer att kunna dra andra fördelar av de nya medierna jämfört med forskare inom naturvetenskap.

Studenter och blivande forskare skolas, mer eller mindre omedvetet, in i olika discipliners kommunikationsmönster. Ett bra tillfälle att skapa nya förhållningssätt gentemot vetenskaplig kommunikation är, enligt Halliday (2003), när studenterna lär sig grunderna i vetenskapligt arbete. Jag håller med Halliday om detta. Oavsett vilken riktning man anser att den vetenskapliga kommunikationen bör ta, kan det vara bra att redan i grundutbildningen diskutera olika modeller inom vetenskaplig kommunikation. Om man ser till vårt eget ämne, Biblioteks- och informationsvetenskap, kan sådan undervisning vara extra viktig. Det antal studenter som går vidare till olika forskningsutbildningar är förhållandevis litet, men åtminstone en del av oss kommer i framtiden att arbeta på högskole- och universitetsbibliotek och där möta forskare med olika behov och inställning till elektronisk egenpublicering.

Jag anser att det är viktigt att föra en diskussion kring olika mediers för- och nackdelar. Det är lätt att förblindas av den nya teknikens alla möjligheter. När förändringar sker bör vi emellanåt stanna upp och fråga oss varför och för vems skull dessa förändringar sker. "Öppna arkiv" som publiceringsmodell är nu några år gammal. Modellen har, som jag ser det, stor potential och skulle kunna leda till helt nya villkor för forskare, universitet, högskolor och bibliotek och kan kanske på sikt leda till ett helt nytt system för vetenskaplig publicering. Men vägen dit tycks lång och många problem återstår att lösa innan detta kan bli verklighet.

6.1. Slutsatser

Som jag nämnde tidigare har mitt val att ha två syften försvårat möjligheterna att på djupet besvara uppsatsen problemformulering och frågeställningar. Jag anser dock att jag utifrån mitt arbete med denna uppsats kan dra följande slutsatser:

- Open Archives Initiative är en återvinningsmodell som rent tekniskt sett fungerar bra, men som kan komma att behöva utvecklas på ett flertal andra områden. En del svagheter med OAI:s återvinningsmodell har uppmärksammats under senare tid, dessa svagheter är till viss del förknippade med mänskliga aspekter av modellen;
- Forskare inom humaniora och samhällsvetenskap har uppenbarligen varit långsammare än många forskare inom naturvetenskap, med att börja arkivera och publicera sina resultat i "öppna arkiv". Men även forskare inom vissa naturvetenskapliga discipliner förhåller sig avvaktande;
- Dessa olikheter när det gäller attityden till "öppna arkiv" har till stor del historiska, kulturella och sociala orsaker;

- Det finns en antydning till att skillnaderna mellan olika discipliner är på väg att utjämnas, i takt med att alla forskare i samma utsträckning får tillgång till datorer och nätverk;
- Olika discipliners behov kommer troligtvis i stor utsträckning avgöra vilken roll "öppna arkiv" kommer att få inom respektive disciplin;
- Biblioteken vid våra högskolor och universitet bör, i uppbyggnaden av "öppna arkiv", vara medvetna om dessa olikheter mellan discipliner;
- Uppbyggnaden av "öppna arkiv" kan komma att leda till nya arbetsuppgifter för dem som arbetar vid universitets- och högskolebibliotek. Biblioteken bör också föra en diskussion om vilka konsekvenser nya publikationsformer kan få för traditionella publikationsformer, t.ex. tryckta monografier. Troligtvis kommer nya samarbetsformer att behöva utvecklas, mellan forskare och biblioteksanställda och mellan olika personalkategorier vid biblioteken;
- Den debatt som förs idag kring vetenskaplig kommunikation domineras, enligt min mening, av ett naturvetenskapligt perspektiv och diskuterar i allt för liten utsträckning hur andra discipliner vill och kan dra fördelar av digitala medier. Jag anser att det humanistiska och samhällsvetenskapliga perspektivet i debatten om den framtida vetenskapliga kommunikationen behöver stärkas.

7. Sammanfattning

Denna uppsats diskuterar "öppna arkiv" (*open archives*), en ny modell för vetenskaplig kommunikation, initierad av organisationen *Open Archives Initiative* (OAI). Denna nya kommunikationsmodell har vuxit fram dels ur ny informationsteknik, dels ur olika forskarkollektivs behov. Uppsatsen har två syften. Det första syftet är att beskriva "öppna arkiv" utifrån ett kunskapsorganisatoriskt perspektiv, det andra syftet är att undersöka "öppna arkiv" utifrån ett socialt, historiskt samt kulturellt perspektiv. Mina två syften preciseras i tre frågeställningar:

- Hur fungerar "öppna arkiv" utifrån ett kunskapsorganisatoriskt perspektiv, d.v.s. vilka metoder används vid lagring, sökning och återvinning av information?
- Vilka eventuella barriärer finns det mot en vidareutveckling av "öppna arkiv" och andra alternativa publiceringsformer?
- Om forskare inom humaniora och samhällsvetenskap i större utsträckning förhåller sig avvaktande till "öppna arkiv" än forskare inom naturvetenskap, vilka tänkbara orsaker kan det i så fall finnas till att det förhåller sig så?

Uppsatsen är skriven utifrån vissa begränsningar. Dessa innebär att jag i uppsatsen huvudsakligen diskuterar "öppna arkiv" utifrån forskarnas och bibliotekens agerande. Den tredje aktören i den vetenskapliga kommunikationskedjan, förlagen, berör jag i mycket begränsad omfattning.

Det finns tre för uppsatsen grundläggande begrepp - kunskapsorganisation, vetenskaplig kommunikation och digitala bibliotek, vilka definieras i inledningskapitlet. Uppsatsens teoretiska ram utgörs av några olika texter av författarna Jack Andersen, Christine L. Borgman samt författarparet Brown och Duguid. Andersen hävdar att ny kommunikationsteknik förändrar vårt sätt att tänka, producera, kommunicera och organisera kunskap och menar att kunskapsorganisation bör sättas in i ett socialt och historiskt sammanhang. Borgman och Brown och Duguid diskuterar även de hur ny informationsteknik påverkar människan sätt att förhålla sig till produktion och kommunikation av kunskap och information. Deras texter fokuserar på samspelet mellan människa och teknik och betonar i hög grad den sociala kontextens betydelse vid skapandet av nya informationsåtervinningssystem.

Den metod som används i uppsatsen är en litteraturstudie. Val av litteratur har sett något olika ut för uppsatsens båda syften. För att besvara uppsatsens första syfte har jag huvudsakligen använt mig av dokument om OAI och dokument publicerade av organisationen OAI och andra aktörer inom alternativ publicering. För att besvara uppsatsens andra syfte har jag använt mig dels av artiklar och monografier om vetenskaplig kommunikation, dels av några olika tidigare genomförda empiriska undersökningar, med anknytning till "öppna arkiv" och vetenskaplig kommunikation.

I kapitel 4 beskriver jag hur "öppna arkiv" fungerar utifrån ett grundläggande kunskapsorganisatoriskt perspektiv, vilket är uppsatsens första syfte. Men kapitel 4 behandlar även delvis uppsatsens andra syfte, att undersöka "öppna arkiv" utifrån ett historiskt, socialt och kulturellt perspektiv, genom att jag där beskriver den vetenskapliga tidskriftens bakgrund samt ger en mer nutida bakgrund till framväxten av OAI. Jag besvarar dock i kapitel 4 huvudsakligen uppsatsens första frågeställning, genom att redogöra för grunderna i OAI:s återvinningsmodell. "Öppna arkiv" är en vidareutveckling av de e-printarkiv som ett antal amerikanska fysiker skapade i början av 1990-talet, arkiv som syftade till ett snabbt och enkelt utbyte av vetenskapliga artiklar i elektronisk form, lagrade i en typ av fulltextdatabaser. Målet med Open Archives Initiative var att hitta de tekniska och organisatoriska lösningarna för samsökning av olika e-printarkiv. OAI:s återvinningsmodell bygger på utbytbara metadataposter enligt formatet Dublin Core, med möjlighet att komplettera med användarspecifika format, samt ett enkelt protokoll kallat *the OAI Protocol for Metadata Harvesting*. Inom OAI:s återvinningsmodell finns tre aktörer. Dataleverantörer (*dataproviders*) kallas de som tillhandahåller arkiv, t.ex. ett universitet. Informationstjänster (*serviceproviders*) samlar in metadataposter från olika dataleverantörer och bygger upp söktjänster anpassade efter den tilltänkta målgruppens behov. Den tredje typen av aktörer är användarna, ofta samma personer som bidrar med material till arkiven.

I kapitel 5 behandlar jag uppsatsen andra syfte, att undersöka "öppna arkiv" utifrån ett historiskt, socialt och kulturellt perspektiv. En viktig utgångspunkt i detta avsnitt är tanken att olika forskarkollektivs val av kommunikationsmedium styrs dels av det eller det medier som står till buds, dels av de olika forskarkollektivens behov. Med hjälp av tidigare genomförda undersökningar och olika texter om alternativa publiceringsmodeller försöker jag finna svar på de två andra av uppsatsens tre frågeställningar: vilka barriärer som kan tänkas förhindra eller försena en ökad användning av "öppna arkiv", samt huruvida humanistiska och samhällsvetenskapliga forskare förhåller sig mer avvaktande till "öppna arkiv" än naturvetenskapligt inriktade

forskare och vad det i så fall kan bero på. Utifrån ett antal olika tänkbara barriärer diskuterar jag hur dessa i allmänhet kan tänkas påverka forskares inställning till egenarkivering och då i synnerhet humanistiskt och samhällsvetenskapligt inriktade forskare. De barriärer jag, i olika utsträckning, diskuterar utifrån är följande: bevarandenaspekter, kvalitetskontroll (peer review), upphovsrätt, val av publiceringsform, tekniska faktorer samt ekonomiska faktorer.

Min studie beskriver i stora drag OAI: s återvinningsmodell och visar på en del av de svagheter som bl.a. några brittiska forskare visat på under senare tid. Modellen tycks tekniskt sett fungera väl. Dock finns ett antal frågor att framöver ta ställning till. En fråga gäller arkivens innehåll. Vilka dokumenttyper och vilka formatter ska tillåtas i arkiven? Vilka metadataformat ska vara tillåtna och vem ska skapa metadata? Dublin Core uppfattas ofta som ett relativt enkelt format att använda. Trots detta har skapandet av metadata genererat problem som bl.a. skulle kunna lösas genom att universitetsbibliotekariernas kunskap tas till vara. En annan fråga gäller metoderna för själva arkiveringen; även här skulle bibliotekspersonal kunna ha en viktig uppgift. Även bevarandenaspekter och olika ekonomiska aspekter behöver i framtiden få större utrymme. Den absolut största frågan enligt vissa är dock hur man ska få forskare att vilja använda arkiven, både genom att bidra med material och genom att använda det material som redan finns arkiverat.

Min studie visar att forskare inom humaniora och samhällsvetenskap har förhållit sig mer avvaktande till alternativa publiceringsmodeller och att det kan finnas historiska, kulturella och sociala orsaker till att det förhåller sig så. Även forskare inom vissa naturvetenskapliga discipliner har dock förhållit sig avvaktande. En del av ovan redovisade barriärer kan tänkas påverka forskare inom alla discipliner, det gäller främst barriärer med anknytning till upphovsrätt och bevarandenaspekter. Barriärer som i större utsträckning kan tänkas påverka forskare inom humaniora och samhällsvetenskap är val av publiceringsform, ekonomiska och tekniska faktorer samt kvalitetsgranskning. "Öppna arkiv" bygger på en kommunikationsmodell som har sitt ursprung inom ett naturvetenskapligt forskarkollektiv och som skapades för att tillmötesgå behov inom just detta kollektiv. Ett av dessa behov var att *snabbt* kunna dela med sig ny kunskap. Detta behov av snabb spridning av resultat anses inte lika stort inom humaniora och samhällsvetenskap, vilket bl.a. visar sig genom att monografin har en mycket starkare ställning inom dessa två forskningsområden, än inom naturvetenskap där tidskriftsartikeln dominerar som publiceringsform. Enligt min uppfattning talas det, i den pågående debatten om "öppna arkiv", betydligt oftare om publicering av tidskriftsartiklar än om monografier, vilket skulle kunna vara en delförklaring till att forskare inom humaniora och samhällsvetenskap inte har anammat den nya modellen i samma utsträckning. En annan förklaring till att humanistiskt och samhällsvetenskapligt inriktade forskare inte varit lika snabba till att börja egenarkivera, kan vara att samhällets mer begränsade tilldelning av ekonomiska bidrag till dessa forskningsområden. Även om "öppna arkiv" sägas vara en förhållandevis billig återvinningsmodell behövs det, åtminstone inledningsvis, ett ekonomiskt utrymme för bl.a. mjukvaruutveckling. Ytterligare en tänkbar faktor är att tillgången till datorer och nätverk har varit mer begränsad inom humaniora och samhällsvetenskap, även om skillnaden jämfört med naturvetenskap är på väg att jämnas ut. När det gäller behovet av traditionell kvalitetsgranskning (peer review) skulle detta kunna vara större inom många humanistiska och samhällsvetenskapliga forskningsområden, som till sin natur ofta är tvärvetenskapliga. När en forskare behöver hämta kunskap från fler områden än

det han eller hon huvudsakligen är verksam inom kan, enligt min åsikt, behovet av traditionell kvalitetsgranskning troligtvis bli större. Än så länge har vissa forskarkollektivs behov av traditionell kvalitetsgranskning inte funnit sin lösning inom modellen med "öppna arkiv".

"Öppna arkiv" är enligt min åsikt en återvinningsmodell med stor potential. Många problem återstår dock att lösa innan "öppna arkiv" kan bli en modell för det stora flertalet forskare. Enligt min mening bör historiska, sociala och kulturella aspekter få större utrymme när nya informationsåtervinningssystem utvecklas, så även i den fortsatta utvecklingen av "öppna arkiv". När det gäller det fortsatta arbetet med "öppna arkiv" bör dessutom det humanistiska och samhällsvetenskapliga perspektivet stärkas.

8. Litteraturlista

Andersen, Jack (2002). Communication technologies and the concept of knowledge organization - a medium-theory perspective. *Knowledge Organization*, 29 (2002) no. 1, s. 29-39.

Anderson, J. D. (1996). Organization of Knowledge. I: J. Feather & P. Sturges (Eds.), *International Encyclopedia of Library and Information Science*, s. 336-353. London & New York: Routledge.

Bacon (1990). I *Nationalencyklopedin*, Bd 2, s. 202-203. Höganäs: Bra Böcker.

Berners-Lee, Tim, Hendler, James & Lassila, Ora (2001). The Semantic Web. *Scientific American*, May 2001, s. 29-37.

Björkhem, Miriam & Lindholm, Jessica (2001). Så funkar det! Metadataformatets roll i det digitala biblioteket. *Tidskrift för dokumentation. The Nordic Journal of Documentation*, vol. 56, no. 1, 2001, s. 1-13.

Borgman, Christine L. (2000a). Digital libraries and the continuum of scholarly communication. *Journal of Documentation*, vol. 56, no. 4, July 2000, s. 412-430.

Borgman, Christine L. (2000b). *From Gutenberg to the global information infrastructure*. Cambridge, Massachusetts: The MIT Press.

Breeding, Marshall (2002a). The emergence of the Open Archives Initiatives. *Information Today*, Apr 2002, vol. 19, Issue 4, s. 46-48.
<http://search.epnet.com/direct.asp?an=6471819&db=afh> [2003-04-15]

Breeding, Marshall (2002b). Understanding the protocol for metadata harvesting of the Open Archives Initiative. *Computers in Library*, Sep 2002, vol. 22 Issue 8, s. 24-30.
<http://search.epnet.com/direct.asp?an=7280939&db=afh> [2003-04-15]

Brown, John Seely och Duguid, Paul (1996). *The Social Life of Documents*. *First Monday*, 1996, May.
<http://www.firstmonday.dk/issues/issue1/documents/index.html> [2002-10-05]

Brown, John Seely & Duguid, Paul (2000). *The social life of information*. Boston: Harvard Business School Press.

Buckland, M. K. (1991). Information as thing. *Journal of the American Society for Information Science*, 42 (5), 1991, s.351-360.

Burke, P. (2000). *A Social History of Knowledge. From Gutenberg to Diderot*. Cambridge: Polity Press.

Cavallin, Mats (2002). e-bokens lov. *Biblis*, no. 18.
http://www.kb.se/biblis/BIBLIS/Artiklar/Nr_18/Cavallin.pdf [2003-01-20]

Costa, Sely & Meadows, Jack (2000). The impact of computer usage on scholarly communication among social scientists. *Journal of Information Sciences*, August 2000, vol. 26, no 4, s. 255-262.

<http://80-www.ingenta.com.lib.costello.pub.hb.se/> [2003-04-15]

Crow, Raym (2002). *The Case for Institutional Repositories: A SPARC Position Paper*. <http://www.arl.org/sparc/IR/ir.html> [2003-03-12]

Deegan, Marilyn & Tanner, Simon (2002). *Digital futures - Strategies for the information age*. London: Library Association Publishing.

Degerstedt, Stina (2002). *Vad är Dublin Core?*

<http://www.kb.se/bus/Metadata/dc/DC:HTM> [2003-04-01]

Dempsey, Lorcan och Heery, Rachel (1997). *Specification for resource description methods. Part 1. A review of metadata: a survey of current resource description formats*.

<http://ukoln.ac.uk/metadata/desire/overview/overview.pdf> [2003-06-09]

Dobratz, Susanne & Matthaei, Birgit (2003). Open Archives activities and experiences i Europe - an overview by the Open Archives Forum. *D-Lib Magazine*, vol. 9, no. 1, January 2003.

<http://www.dlib.org/dlib/january03/dobratz/01dobratz.html> [2003-02-12]

Finneman, N. O. (1999). Modernity Modernised - The Cultural Impact of Computerisation. I.: Paul A. Mayer (ed.) *Computer Media and Communication - A Reader*, s.141-160. Oxford University Press.

<http://www.hu.au.dk/ckulturf/pages/publications/nof/mom.pdf> [2003-06-05]

Gadd, Elizabeth, Oppenheim, Charles & Proberts, Steve (2003). *RoMEO Studies 1: The impact of copyright ownership on academic author self-archiving*

[http://www.lboro.ac.uk/departments/ls/disresearch/romeo/RoMEO Studies 1.pdf](http://www.lboro.ac.uk/departments/ls/disresearch/romeo/RoMEO%20Studies%201.pdf) [2003-04-15]

Greenberg, Jane, Pattuelli, Maria Christina, Parsia, Bijan & Davenport Robertson, Dave (2001). Author-generated Dublin Core metadata for web resources: A baseline study in an organization. *Journal of Digital Information*, vol. 2, issue 2.

<http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Greenberg/> [2003-04-15]

Greenberg, Jane & Davenport Robertson, Dave (2002). Semantic Web construction: An inquiry of authors' views on collaborative metadata generation. *DC-2002: Metadata for e-Communities: Supporting Diversity and Convergence. Proceedings for the International Conference on Dublin Core and Metadata for e-Communities, 2002, Florence, Italy. October 13-17. Firenze*. University Press, s. 45-52.

<http://www.bncf.net/dc2002/program/ft/paper5.pdf> [2003-01-20]

Guédon, Jean-Claude (2001). In Oldenburg's long shadow: librarians, research scientists, publishers, and the control of scientific publishing. *ARL Proceeding 138 May 2001, Membership Meeting: Creating the digital future*.

<http://www.arl.org/arl/proceedings/138/grdon.html> [2003-02-07]

- Hagerlid, Jan (2002). Systemskifte på väg inom vetenskaplig publicering - nu vill forskarna återta kontrollen. *Tidskrift för dokumentation. The Nordic Journal of Documentation*, 57 (2002) 3, s. 95-106.
Även: <http://www.lib.chalmers.se/cthb/Hagerlid1.pdf> [2003-04-30]
- Halliday, Leah (2001). Scholarly communication, scholarly publication and the status of emerging formats. *Information research*, vol. 6, No 4, July 2001.
<http://informationr.net/ir/6-4/paper111.html> [2003-03-25]
- Hansson, Joakim (1999). *Klassifikation, bibliotek och samhälle: En kritisk hermaneutisk studie av "Klassifikationssystem för Svenska Bibliotek"*. (Skrifter från Valfrid, 19). Borås: Valfrid.Diss. Göteborgs universitet.
- Harnad, Steven (2001). *Nature debates: The self-archiving initiative*.
<http://www.nature.com/nature/debates/e-access/Articles/harnad.html> [2003-01-24]
- Hjörland, B. (1997). *Information Seeking and Subject Representation: An Activity-Theoretical Approach to Information Science*. Westport, Connecticut, London: Greenwood Press.
- Hjörland, B. (1998). The classification of psychology: a case study in the classification of a knowledge field. *Knowledge Organization*, 24 (4), 162-201.
- Hunt, Cathy (2002). *Archive User Survey*.
<http://www.eprints.org/results/>
- Hurd, Julie M. (2000). The transformation of scientific communication: A model for 2020. *Journal of the American Society for Information Science*. 51 (14): 1279-1283, 2000.
http://80-www3.interscience.wiley.com.lib.costello.pub.hb.se/cgi-bin/fulltext/74500450/FILE?TPL=ftx_start [2003-04-15]
- Johnson, Richard K. (2002). Institutional repositories. Partnering with faculty to enhance scholarly communication. *D-Lib Magazine*, November 2002, vol. 8, no. 11.
<http://www.dlib.org/dlib/november02/johnson/11johnson.html> [2003-04-15]
- Kjellberg (2001). *Rapport för kartläggning gällande interoperabilitet mellan LUs olika metadataformat och samlingar*.
http://www.lub.lu.se/~sara/digbib/kart_rapport_2001-11-30.html [2003-01-22]
- Kling, Rob och McKim, Geoffrey (1999). Scholarly communication and the continuum of electronic publishing. *Journal of the American Society for Information Science*, 50 (10): 890-906.
<http://80-download.interscience.wiley.com.lib.costello.pub.hb.se/cgi-bin/fulltext?ID=62502159&PLACEBO=IE.pdf> [2003-04-15]

Kling, Rob och McKim, Geoffrey (2000). Not just a matter of time: Field differences and the shaping of electronic media in supporting scientific communication. *Journal of the American Society for Information Science*, 51 (14):1306-1320.
http://80-www3.interscience.wiley.com.lib.costello.pub.hb.se/cgi-bin/fulltext/75000719/FILE?TPL=ftx_start [2003-04-15]

Koehler, Wallace (2002). A call to action: What every searcher should know--and do--about domain names, standards and metadata. *Searcher*, Oct 2002, vol. 10, Issue 9.
<http://lib.costello.pub.hb.se/menu> [2003-01-22]

Lagoze, Carl & Van de Sompel, Herbert (2001). The Open Archives Initiative: building a low-barrier interoperability framework. *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries, Roanoke VA, 17-23 juni. 2001*, s. 54-62
www.openarchives.org/documents/oai.pdf [2003-01-20]

Lagoze, Carl, Van de Sompel, Herbert, Nelson, Michael & Warner, Simeon (2002). *The Open Archives Initiatives Protocol for Metadata Harvesting*.
www.openarchives.org/OAI/2.0/openarchivesprotocol.html [2003-03-17]

Lindholm-Romantschuk, Ylva (1996). The role of monographs in scholarly communication: an empirical study of philosophy, sociology and economics. *The Journal of Documentation*, vol. 52, no 4, December 1996, s. 389-404.

Lindholm-Romantschuk, Ylva (1998). *Scholarly book reviewing in the social sciences and humanities. The flow of ideas within and among disciplines*. Westport: Greenwood Press.

Lundberg, Sigfrid & Lindholm, Jessica (2003). *Samordning av söktjänster och metadata-repositorier för examensarbeten*.
<http://www.lub.lu.se/lucep/activities//oai-etc/> [2003-04-24]

Meadows, Arthur Jack (1974). *Communication in science*. London: Butterworth.

Meadows, Arthur Jack (1998). *Communicating research*. London: Academic Press.

Odlyzko, Andrew (2002). The rapid evolution of scholarly communication. *Learned Publishing*, (2002) 15, s. 7-19.
<http://www.catchword.com/alpsp/09531513/v15n1/contp1-1.htm> [2003-04-15]

Owen, John Mackenzie (2000). From hybrid to virtual scientific communication: implications for the information professional. *Paper presented at the New Book Economy – BIS Symposium, Council of Europe, Strasbourg, 3-4 February 2000*.
<http://cf.hum.uva.nl/bai/home/jmackenzie/pubs/Straatsburg.htm> [2003-01-22]

Pinfield, Steven, Gardner, Mike & MacColl, John (2002). Setting up an institutional e-print archive. *Ariadne*, Issue 31, April 2002.
<http://www.ariadne.ac.uk/issue31/eprint-archives/> [2003-04-15]

- Pinfield, Steve (2003). Open Archives and UK Institutions. An Overview. *D-Lib Magazine*, vol. 9, no. 3, 2003.
<http://www.dlib.org/dlib/march03/pinfield/03pinfield.html> [2003-04-15]
- Rabow, Ingegerd (2001). Ideals over deals? Den vetenskapliga kommunikationsmarknaden *Tidskrift för dokumentation. The Nordic Journal of Documentation*, 56 (2001) 2, s. 33-46.
- Rodman Annika (2003). *Informationskompetens: Självstudier på nätet*.
<http://www.lub.lu.se/ub/distans/infokompetens/vetkomm/vetkomma.html> [2003-04-20]
- Rowland, Fytton (1997). Print journals: Fit for the future? *Ariadne*, Issue 7, January 1997.
<http://www.ariadne.ac.uk/issue7/fytton/> [2003-02-19]
- Rusch-Feja, Diann (2002). The Open Archives Initiatives and the OAI protocol for Metadata Harvesting: rapidly forming a new tier in the scholarly communication infrastructure. *Learned Publishing*, vol. 15, no. 3, July 2002.
<http://www.catchword.com/alpsp/09531513/v15n3/contp1-1.htm> [2003-02-25]
- Savenije, Bas (2002). Who pays the ferryman? About new models for scientific communication. *Paper presented at the international symposium "Science & engineering libraries for the 21st century", Leuven, Belgium, October 2-4, 2002*, s. 1-13.
http://www.figaro-europe.net/docs/BAS_Leuven.doc [2003-04-15]
- Savenije, Bas & Smith, Jennifer (2002). *Conference Paper: El Pub 2002: The FIGARO Project: Realising New Models in Scientific Communication*
http://www.figaro-europe.org/index2.html?about_us.html [2003-01-22]
- scholarship (1980). I *Stora engelsk-svenska ordboken*, s. 778. Nacka: Esselte Studium.
- science (1980). I *Stora engelsk-svenska ordboken*, s. 778. Nacka: Esselte Studium.
- Smith, Alastair G. (2002). Does metadata count? A webometric investigation. *DC-2002: Metadata for e-Communities: Supporting Diversity and Convergence. Proceedings for the International Conference on Dublin Core and Metadata for e-Communities, 2002, Florence, Italy. October 13-17. Firenze*. University Press, s. 133-138.
<http://www.bncf.net/dc2002/program/ft/paper15.pdf> [2003-01-10]
- Solomon, David.J (2002). Talking past each other: Making sense of the debate over electronic publication. *First Monday*, vol. 7, no. 8, August 2002, s. 29-39.
http://www.firstmonday.dk/issues/issue7_8/solomon/index.html [2003-01-24]
- Suber, Peter (2003). *Removing the barriers to research: an introduction to open access for librarians*.
<http://www.earlham.edu/~peters/writing/acrl.htm> [2003-04-15]

Swan, A. (2002). *Authors and electronic publishing: the ALPSP research study on authors' and readers' views of electronic research communication*. West Sussex, ALPSP: 83pp.

The Knight Higher Education Collaborative (2002). *Op.cit.: publishing in the humanities and social sciences*.

Learned Publishing (2002) 15, no. 3, July, s. 205-216.

<http://www.alpsp.org/volcont.htm> [2003-04-15]

Tomaiuolo, Nicholas G. & Packer, Joan G. (2000). Preprint servers: pushing the envelope of electronic scholarly publishing. *Searcher*, 10 704 795, vol. 8, no. 9, October 2000.

<http://www.infoday.com/searcher/oct00/tomaiuolo&packer.htm> [2003-02-05]

White Howard D. & McCain Katherine W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972-1995. *Journal of the American Society for Information Science*, 49 (4) s.327-355.

Webbplatser

Arkiv Ex –Sverige (2003).

<http://sigge.lub.lu.se/phptest/2003/arkiv-ex/> [2003-04-26]

arXiv.org e-Print archive (2003).

<http://arxiv.org/> [2003-04-15]

Budapest Open Access Initiative (2002). *BOAI homepage*.

<http://www.soros.org/openaccess/read.shtml> [2003-03-05]

Cogprints (2003).

<http://cogprints.ecs.soton.ac.uk/> [2003-04-15]

E-LIS (2003).

<http://eprints.rclis.org/> [2003-04-15]

eprints.org (2003)

<http://www.eprints.org> [2003-06-09]

FIGARO (2003). *FIGARO homepage*.

<http://www.figaro-europe.net> [2003-01-20]

my.OAI (2003).

<http://www.myoai.com/> [2003-04-15]

Open Archives Initiative (2003). OAI homepage.

<http://www.openarchives.org> [2003-04-26]

perseus.tufts.edu (2003).

<http://www.perseus.tufts.edu> [2003-03-17]

Public Library of Science (2003). *PloS Homepage*.
<http://www.publiclibraryofscience.org/people.htm> [2003-03-05]

Scholarly Publishing and Academic Resources Coalition (2003). *SPARC homepage*.
www.arl.org/sparc/ [2003-03-25]