Om allmänna handlingars autenticitet vid digitalt långtidsbevarande

Catharina Grönqvist & Jonas Hult

Examensarbete (30 högskolepoäng) i arkivvetenskap för masterexamen inom ABM-masterprogrammet vid Lunds universitet.
Handledare: Isto Huvila
År: 2013

Abstract

Title – "On the authenticity of public records in long term digital preservation"

Purpose –The documenting and administration of society is becoming increasingly digital. In analogue preservation the archival material´s physical properties is instrumental in determining authenticity. Digital material on the other hand, is more dependent on intellectual control and metadata in order to maintain authenticity. It is the aim of this thesis to examine in which ways this digital authenticity could be maintained and to apply this knowledge in an analysis of three of the larger public Swedish archival institutions. The thesis has a user perspective connected to the other main question, which is to analyze which possibilities an archives user in the distant future might have to determine the authenticity of the digital public records created today.

Methodology/approach – The mapping of the methods and theories concerning the maintenance of digital authenticity are pursued through an interdisciplinary literature study of the subjects archival science, digital diplomatics and digital preservation. An empirical study was carried out concerning the Swedish National Archives, Stockholm Municipal Archives and SYLL, a joint university project, to establish which methods were used in their digital preservation strategies. The results were analyzed and sorted with the use of systems theory, the OAIS-model and diplomatics. The latter suitable for interdisciplinary studies both qualitative and quantitative in character.

Findings – Documentation concerning the authenticity of the digital objects in the public archives is created according to *best practice*. That means that authenticity is being preserved using the latest and most suitable technical methods and theories. However, some methods concerning authenticity believed to be important in archival science and diplomatics a decade ago – such as digital signatures – are not used today because deemed impractical. Owing to the chosen preservation strategy in Sweden, migrations of the digital material will become inevitable, creating a need for new kinds of authenticity markers. A problem specific for Sweden is that the archives cannot exert control over migration processes before the digital documents reach the archives. The independent standing of governmental agencies makes it difficult to dictate exactly how they should manage documents in their care, the consequence being that the final delivery to the archives creates an inflection point concerning the possibility to ensure authenticity. The future user will however have a number of possibilities to argue for the authenticity of the public record on his computer screen. Detailed metadata, documentation and technical methods, the best practice of which will be guaranteed through the principle of the neutrality of the third party, result in good conditions. Finally, archives having the characteristics of closed systems will probably continue to facilitate public trust in the safe keeping and authenticity of the digital documents.

Originality/Value – An empirical study of how digital authenticity is actually maintained in public archival institutions has not been carried out in Sweden. North American studies do exist, but without the user perspective and unfortunately with limited results owing to linguistic confusion between researcher and informants. The

interdisciplinary collection of digital authenticity markers compiled in this study is unique and can be used as a basis for further research.

Keywords – digital; preservation; authenticity; diplomatics; systems theory; OAIS; archival science; ALM; archive; trust; trusted third-party; records; digitalt bevarande; autenticitet; diplomatik; systemteori; arkivvetenskap; ABM; tredjepartsintygande; allmänna handlingar; arkiv; tillit

Innehåll

1. Inledning 6
1.1 Saker är inte alltid vad de verkar 6
1.2 Uppsatsens disposition 7
2. Bakgrund 8
2.1 Arkivens samhällsroll 8
2.2 Världen digitaliseras 9
2.3 Autenticitet och det digitala samhället 13
3. Forskningsproblem 14
4. Syfte, frågeställningar och avgränsningar 15
4.1 Syfte 15
4.2 Frågeställningar 15
4.3 Avgränsningar 16
4.4 Definitionsdiskussion 17
4.5 Autenticitet och tillit 21
5. Tidigare forskning 23
5.1 Vad bevaras i e-arkivet? 23
5.2 Bevara eller lagra – autenticitet inom digitalt bevarande 26
6. Metod 30
6.1 Övergripande metod 30
6.2 Metodik gällande första delfrågeställningen 31
6.3 Metodik gällande övriga delfrågeställningar 32
6.3.1 Institutioner och informanter – presentation och urval 34
6.4 Forskningsetik 34
6.5 Resultatbearbetning och analys 36
7. Teoretiska perspektiv 38
7.1 Systemteori 39
7.2 Den infologiska ekvationen 41
7.3 Open Archival Information System – OAIS 42
7.4 Diplomatiken 50
7.4.1 Diplomatikens digitala anpassning 51
7.4.2 Integritet 52
7.4.3 Proveniens 53
7.5 Tredjepartsintygandet – institutionen som neutral garant 54
7.6 Det teoretiska perspektivets påverkan och användning 54
8. Resultat 57
8.1 Hur bevaras autenticitet? 57
8.1.1 Digitalt bevarande 57
8.1.2 Autenticitetsmarkörer 59
8.1.3 Signifikanta egenskaper och kopior 62
8.1.4 Teoretisk belysning av litteraturöversikten 65
8.2 Information eller handling? 66
8.3 Metoder och autenticitet 67

8.3.1 Autenticitetsmarkörer som används av institutionerna 67
8.3.2 De viktigaste autenticitetsmarkörerna 72
8.4 Vilka kommer användarna att vara? 73
8.5 Överlämningen till e–arkivet 74
9. Analys 77
9.1 Systemteoretisk återkoppling 77
9.2 Bevarandestrategin påverkar systemet 80
9.3 Varför har vissa autenticitetsmarkörer valts ut? 81
9.4 Institutionernas val och de framtida användarna 85
10. Diskussion 90
10.1 Resultatdiskussion 90
10.2 Metoddiskussion 95
11. Slutsats 99
12. Referenslista 101
Bilaga 1: Intervjufrågor 109
Bilaga 2: Presentationsbrev till institutionerna 110
Bilaga 3: Arbetsfördelning 111

1. Inledning

1.1 Saker är inte alltid vad de verkar

Mellan 2000 och 2005 publicerade den brittiske historikern Martin Allen tre verk som totalt ändrade bilden av britternas agerande under andra världskriget.[1] Baserat på dokument som Allen tagit del av på det brittiska nationalarkivet kunde han avslöja att självaste hertigen av Windsor, senare krönt till kung Edward VIII, under kriget fört vidare statshemligheter till Tyskland och att den brittiske premiärministern hade bedrivit hemliga förhandlingar med Tyskland bakom ryggen på sina allierade. Vidare hade Heinrich Himmler enligt Allen inte alls begått självmord, utan hade i själva verket blivit mördad av brittiska agenter på Winston Churchills order, antagligen för att dölja Churchills förräderi. Dessa uppgifter var ett kraftigt angrepp på den historiska brittiska självbilden av totalt motstånd mot det nazistiska Tyskland (Katz 2008). Det stora problemet var att samtliga dokument som avslöjandena var hämtade från var förfalskningar. Redan 2004 hade misstankar uppstått att de dokument Allen använde som källor och som förvarades i det brittiska nationalarkivets serier i själva verket var inautentiska. Mellan 2005 och 2007 hittades i arkivet 29 förfalskade dokument, av vilka samtliga hade använts av Martin Allen som källor (The National Archives uå). Det som avslöjade dokumenten som förfalskningar var dels att den fysiska informationsbäraren – pappret – i en forensisk undersökning konstaterades inte tillhöra perioden, dels en kontextuell undersökning av informationsinnehållet som avslöjade felaktiga titlar på omnämnda personer och anakronistiskt användande av vissa uttryck (Katz 2008). De falska dokumenten hade smugglats in i serierna för att senare "upptäckas" av Allen i hans forskning.

1998 blev den amerikanske forskaren Thomas Lowry känd för sin upptäckt av ett dokument i det amerikanska nationalarkivet som visade att president Abraham Lincolns sista handling innan han begav sig iväg till teatern där han mördades den 14 april 1865, var att benåda den unge desertören Patrick Murphy från ett dödsstraff.[2] Händelsen blev snabbt en vida spridd anekdot som berättades om den populäre Lincoln. Men även detta dokument visade sig vara inautentiskt och en förfalskning. Efter att arkivarien Trevor Plante lagt märke till att siffran "5" såg annorlunda ut i dokumentets kontext lyckades nationalarkivet till sist få Lowry att 2011 erkänna att

[1] De tre verken var 1. *Hidden Agenda: How the Duke of Windsor Betrayed the Allies* (London, Macmillan, 2000). 2. *The Hitler / Hess Deception: British Intelligence's Best Kept Secret of the Second World War* (HarperCollins, London, 2003). 3. *Himmler's Secret War: the Covert Peace Negotiations of Heinrich Himmler* (Robson Books, London, 2005).

[2] Publicerad i *Don't Shoot That Boy: Abraham Lincoln and Military Justice*, 1999.

han smugglat in en penna till arkivet för att i dokumentet ändra det ursprungliga "1864" till "1865" (National Archives and Records Administration, NARA 2011). Med detta ändrades dokumentets historiska betydelse kraftigt då benådandet inte var presidentens sista handling utan skedde ett år före hans död.

Det första exemplet visade hur falska dokument stoppas in i befintliga serier, medan det andra handlade om hur autentiska dokument kan korrumperas genom obehörig åtkomst. I det första fallet var ursprunget – proveniensen – felaktigt, i det andra var dokumentets integritet bruten. I båda fallen var dock mediets fysiska karaktär en viktig nyckel till att avslöja inautenticiteten. Om Lincoln och Churchill hade använt ett digitalt medium för korrespondens och för att dokumentera sina aktiviteter hade fallen med säkerhet sett annorlunda ut. För framtidens historieskrivare, och användare överhuvudtaget, kommer källmaterialets digitala karaktär i många fall vara en realitet. Andra metoder måste till när tilltron till arkivmaterial bestående av ettor och nollor ska bedömas. Det är denna digitala autenticitet följande undersökning kommer att behandla.

1.2 Uppsatsens disposition

Uppsatsens bakgrundskapitel är tänkt att skissera sammanhanget för uppsatsens forskningsområde, därpå ringas forskningsproblemet in och syftet med undersökningen formuleras. Detta utmynnar i de exakta frågeställningar som ska besvaras. För att kunna besvara dessa diskuteras och definieras autenticitetsbegreppet, varpå ett kapitel över tidigare forskning rörande autenticitet inom digitalt bevarande tar vid. Härefter tar metodkapitlet vid vilket beskriver hur undersökningen bedrivits och vilka avvägningar som gjorts kring detta. Därpå presenteras de teoretiska perspektiv som påverkat undersökningen och tolkningen och analysen av dess resultat. Resultatkapitlet besvarar därefter de delfrågeställningar som formulerats med hjälp av de två forskningsmetoder som undersökningen grundar sig på, litteraturgranskningen och intervjustudien. Forskningsfrågan besvaras i analyskapitlet och dess konsekvenser och vidare samhällsrelevans diskuteras därefter i diskussionen. Uppsatsen avslutas med vilka slutsatser som kan dras från undersökningen.

2. Bakgrund

2.1 Arkivens samhällsroll

Vad innebär då begreppet "arkiv", ett ord som de flesta människor någon gång använder? För ungefär hundra år sedan kodifierade en grupp holländska arkivarier den existerande arkivteorin i Tyskland och Frankrike. Deras tes om vad som utgör ett arkiv kan formuleras såhär på engelska: "The whole of the written documents, drawings and printed matter, officially received or produced by an administrative body or one of its officials" (Muller, Feith & Fruin 1968 i Hirtle 2000, s. 10). Genom arkivprocesser förses dokumenten med kontext, trovärdighet och autenticitet, alla nödvändiga egenskaper. Hirtle (2000, s. 10) menar vidare att det är just egenskapen att informationen är kontextuell som är avgörande för arkivbegreppet. Det är alltså inte tillräckligt med en blandning av all möjlig information utan några särskilda inbördes kopplingar eller beskrivningar.

För de offentliga arkiven regleras verksamheten av lagar. En av dessa är arkivlagen (1990:782) vilkens tredje paragraf lyder att myndigheternas arkiv skall "bevaras, hållas ordnade och vårdas så att de tillgodoser
1. rätten att ta del av allmänna handlingar,
2. behovet av information för rättskipningen och förvaltningen, och
3. forskningens behov."

I samma paragraf betonas att myndigheternas arkiv är en del av det nationella kulturarvet. Sålunda blir de tänkta brukarna tydliga – arkiven skall tjäna allmänheten, statens tjänstemän och forskarna. När det gäller allmänheten är offentlighetsprincipen och rätten att ta del av allmän handling bärande. Offentlighetsprincipen ingår i grundlagen tryckfrihetsförordningen (1949:105). Här fastställs att allmänna handlingar, med vissa inskränkningar som beskrivs i offentlighets- och sekretesslagen (2009:400), skall vara offentliga. Bohlin (2010, s. 22) poängterar att medborgarnas rätt att ta del av allmänna handlingar är en av de grundläggande betingelserna för en fri och demokratisk åsiktsbildning. Omvänt kan det även sägas vara ett skydd för myndigheterna som genom denna insyn kan skydda sig mot anklagelser om maktmissbruk, korruption och oegentligheter. Det grundläggande begreppet "allmän handling" behöver dock en definition för att vara användbart. En handling är allmän om den är inkommen eller upprättad på en myndighet och förvaras på myndigheten[3] (Bohlin 2010, ss. 40ff). Inom myndigheter motsvaras begreppet allmän handling

[3] 2 kap. 3§ tryckfrihetsförordningen (1949:105).

överlag av begreppet arkivhandling (Gränström, Lundqvist & Fredriksson 2000, s. 16). Hur de allmänna handlingarna relaterar till andra arkivvetenskapliga termer såsom dokument och *records* och även termer inom digitalt långtidsbevarande som digitalt objekt kommer att diskuteras djupare under avsnitt 5 om tidigare forskning. För tillfället räcker det dock att konstatera att de offentliga arkivmyndigheterna har en skyldighet enligt arkivlagen att allmänna handlingar skall "bevaras, hållas ordnade och vårdas". I sjätte paragrafen i samma lag specificeras vad som skall ingå i arkivvården, bland annat skall myndigheten "organisera arkivet på ett sådant sätt att rätten att ta del av allmänna handlingar underlättas", "skydda arkivet mot förstörelse, skada, tillgrepp och obehörig åtkomst" och "verkställa föreskriven gallring i arkivet." Just dessa tre aspekter av arkivvården kan komma att kompliceras då de digitala arkiv som nu levereras ställer nya krav på bevarandet och har andra förutsättningar än de äldre, analoga pappersarkiven. Några av dessa förutsättningar kommer att presenteras i följande avsnitt.

2.2 Världen digitaliseras

Allt fler av världens transaktioner och informationsskapande sker digitalt. Tidningsläsning sker via Internet, ekonomiska transaktioner sätts aldrig på pränt och en stor del av all korrespondens sker via e-post, för att nämna några exempel. 2008 skapades 1.5 exabyte[4] unik information, vilket var samma mängd information som i nedtecknad form genererats av mänskligheten från civilisationens vagga till år 2000 (Ferguson 2008). Samson uppgav 2009 (s. 179) att mängden digital information dubbleras var 18:e månad, medan uppgifter från oktober 2012 menade att under 2012 skapades 2.5 exabyte information *varje dag* och att denna mängd dubbleras var fyrtionde månad (McAfee & Brynjolfsson 2012). Det ter sig alltså som om att mängden information som existerar just nu är svår att definiera eftersom skapandet av ny information har en ökningshastighet som är svår att förutse över längre tidsrymder. Det rör sig oavsett vilket om enorma mängder information, men den absoluta lejonparten av denna kommer inte att bevaras av arkiv i framtiden, delvis på grund av att den skapas och delas av privatpersoner på Internet vilket medför frågetecken kring ansvaret för bevarandet. Att välja ut vilken digital information som ska bevaras är en fråga som diskuteras inte minst inom arkivvetenskapen, vilket fångas av Gladney i hans översatta uttalande att "om all den här informationen var värd att skapas måste en del av den vara värd att spara" (2007, s. 3). Man kan åtminstone konstatera att på det nationella planet är det absolut nödvändigt att kunna bevara handlingarna som skapas digitalt inom offentlig förvaltning, exempelvis sjukvårdsjournaler, födelsebevis och lagfarter, under många år för att samhället ska kunna fungera. Ett exempel kan illustrera hur viktigt det är med ett välfungerande system för e-arkivering: små skillnader i arkiveringspraxis och informationsbevarande och -tillgängliggörande när det gäller pensionsgrundande uppgifter kan få stora ekonomiska konsekvenser för den enskilde medborgaren (Askergren 2009, ss. 187ff).

[4] Prefixet exa står för 10^{18}, alltså 1 000 000 000 000 000 000.

Den snabba tekniska utvecklingen och digitaliseringen av de dokument vi människor lämnar efter oss är en av de mest diskuterade frågorna inom arkivvärlden. Två huvudspår i diskussionen kan skönjas, dels behandlar den rent tekniska frågor om bevarande och lagringsmedia, dels försöker man inlemma arkivteoretiska resonemang i samband med den digitala miljön. Det är viktigt att även göra en distinktion här gällande elektroniska dokument. Å ena sidan finns dokument som är *born digital*, de är skapade digitalt och sparas digitalt. Å andra sidan finns analogt material som digitaliserats. Detta görs av tillgänglighetsskäl, utrymmesskäl eller för att minska slitaget på originalen. Oavsett om det digitala arkivmaterialet är digitaliserat eller *born digital* måste det beskrivas med information i form av metadata för att möjliggöra bevarande, återsökning och användning (Taylor & Joudrey 2009, ss. 89ff). Nedanstående figur visar hur metadata kan kategoriseras i tre övergripande kategorier som syftar till att dokumentera olika aspekter hos digitala objekt.[5]

Figur 1. Tabell över metadatakategorier

Metadatakategori	Syfte eller funktion	Dokumenterar	Källa
Administrativ	1. Behövs för att datorn ska kunna återge det digitala objektet. 2. Ger information om tekniska aspekter. 3. Information om bevarandet	1. Filformat, relevant programvara. 2. Storlek, born-digital/digitaliserad. 3. Övertagande från annat arkiv, säkerhetsåtgärder, migreringar.	Taylor & Joudrey 2009, ss. 96-98.
Strukturell	Förklarar relationer mellan mindre beståndsdelar i sammansatta digitala objekt, därför överflödig för enstaka objekt.	Beskriver hur tabeller i databaser eller delarna och funktionerna i en hemsida relaterar till varandra.	Taylor & Joudrey 2009, ss. 100-102.
Deskriptiv	Beskriver informationen i det digitala objektet.	Arkivbildare, skapandekontext, skapandetid, ämne.	Taylor & Joudrey 2009, ss. 103ff.

Sammanställning gjord av författarna

Metadata av teknisk natur är en förutsättning för att digitala objekt ska kunna läsas av mjukvara samtidigt som kontextskapande metadata krävs för att objektet ska kunna förstås av användare. Detaljnivån för metadatan hos digitala objekt kan variera då ett flertal faktorer kan påverka, bland annat huruvida metadatan genereras automatiskt av mjukvara eller om mänsklig inblandning behövs. Automatisk metadata kan naturligtvis inte beskriva alla deskriptiva aspekter hos ett digitalt objekt utan är främst anpassat för att dokumentera administrativa faktorer. Människor behövs exempelvis fortfarande för att göra explicita tolkningar och för att sätta arkivet i sin kulturella kontext (Frendo 2007, s. 167).

Tiden är en viktig faktor för vilken metadata som bör läggas till ett digitalt objekt eftersom man i ett längre tidsperspektiv endast kommer att ha den information som finns lagrad i metadatan som en resurs för att göra objektet läsbart och förståeligt. Det som idag uppfattas som självklart och är en tyst kunskap, till exempel att vissa format

[5] Termen digitalt objekt kan användas för att beteckna en digital arkivhandling, ett digitalt dokument i form av en PDF-fil, en bild och så vidare. Begreppet används gärna inom digitalt bevarande eftersom det är tämligen inkluderande. En djupare diskussion och definition kring termen återfinns i avsnittet 5.1.

innebär att de är skapade av vissa program och att dessa program behövs för att öppna dem, kommer nödvändigtvis inte att vara lika självklart för användare om 50 eller 100 år.

Hur ett digitalt objekt beskrivs med metadata, vilka aspekter av objektet som beskrivs och indexeras och även, i vissa fall, hur dessa beskrivningar formuleras, påverkar den framtida överlevnaden, användbarheten och tillgängligheten hos ett digitalt dokument. Metadata som rör förutsättningar för bevarande och användande kommer att behandlas utförligare i uppsatsens teoridel medan det här kan konstateras att den deskriptiva metadata som främst används för återsökning till stor del påminner om den "metadata" som hört till traditionella analoga arkiv, så som arkivbeskrivningar och förteckningar. Denna typ av metadata är mer subjektiv till sin natur än de administrativa och strukturella metadata som hör samman med bevarande och användande (Ibid., ss. 92, 102–103). På grund av den deskriptiva metadatans subjektiva karaktär påverkar de formuleringar som arkivarien väljer återsökningsmöjligheterna för framtida användare, potentiellt både till det bättre och det sämre (Duff & Harris 2002, ss. 81–82). Även utformningen av det elektroniska sökgränssnittet gentemot arkivmaterialet påverkar möjligheterna att återfinna materialet (Ibid., s. 80). Hur sökgränssnitten utformas påverkar dock vilka förväntningar användarna får på sökningen och materialet som genereras i denna (Johnson 2008, s. 161); påminner ett sökverktyg inom ett arkiv för mycket om gränssnittet på exempelvis Google kan det skapa föreställningar hos användaren om att sökningen gått till på samma sätt och att resultaten presenteras likartat, vilket inte behöver vara fallet.

En aspekt när det gäller de samtida användarna beskrivs av Cox (2007). Teknikutvecklingen har gett användarna väldigt höga förväntningar i mötet med arkiven och dess personal. Det gäller även, menar Cox, förväntningar på att hela samlingar ska vara, eller snabbt bli, digitaliserade. Ilshammar (2008) utökar resonemanget med att diskutera förhållandet mellan arkivarien och uppdragsgivaren i samband med digitalisering. Ofta finns från politiskt håll en övertro på tekniken och man kan höra resonemang om att digitalt material inte tar någon plats och att det som digitaliserats håller i långa tidsperioder. Så är det inte. Förutom värdering, kontextualisering och den bearbetning med den för digitaliserat material så viktiga metadatan, finns även problem med bristande resurser och den olösta tekniska frågan med långtidslagring (Ilshammar 2008). Conway (2010) menar även han att digitaliseringen mest från uppdragsgivarroll ses som en tillgänglighetsstrategi, men i lagringen finns reella problem som fysiskt sönderfall och digital obsolescens.

Papper, om det är av god kvalitet och lagras under gynnsamma omständigheter, kan bevaras under hundratals år utan att det faller sönder eller degenererar till den grad att informationsinnehållet inte längre är läsbart. Digitala lagringsmedia, eller databärare, är dock inte riktigt lika hållbara, vilket tabellen nedan visar.

Figur 2. Tabell över olika lagringsmedias hållbarhet

Lagringsmedia	**Exempel**	**Beräknad hållbarhetstid i år**
Hårddisk	Dator	3 – 6
Magnetband	VHS-kassetter	10 – 20
Magnetdisketter	3.5 tums disketter	1–5
Optiska diskar	CD, DVD	10 –100
Statiskt minne	USB, flashminnen	50 – 100

ZDNet Australien

Detta medför att nya utmaningar tillkommer i elektroniska arkiv då databärarna måste förnyas med, i jämförelse med analoga arkiv, tämligen korta tidsintervall.

Förutom att lagringsmedia sönderfaller uppstår problem i det digitala bevarandet genom att det ständigt tas fram nya versioner av den mjukvara som bland annat används för att läsa eller skapa filer i vissa format. Dessa nya mjukvaruversioner är inte alltid kompatibla med äldre filformat vilket gör att dessa då inte kan användas. För att åstadkomma ett digitalt bevarande finns tre huvudsakliga strategier: migrering, emulering eller museimetoden. *Museimetoden* grundar sig i att man försöker bevara så ursprungligt som möjligt med datorer, servrar, databärare och så vidare som är så lika och samtida i konstruktion som de de ursprungliga datafilerna skapades på. *Emuleringsstrategin* består översiktligt i att den programvara som användes för att skapa en fil bevaras och filen öppnas i det programmet (Galloway 2004, s. 574; Gladney 2007, ss. 238ff; Giaretta 2011, ss. 197ff). Det kan dock krävas att andra program måste tas fram för att den bevarade programvaran ska kunna användas i nyare operativsystem eftersom dessa inte nödvändigtvis är kompatibla med gammal programvara. Den sista, och mest utbredda strategin, är *migreringsstrategin* vilken syftar till att uppdatera filformaten som filerna finns i till nyare versioner eller andra mer arkivvänliga filformat (Galloway 2004, ss. 574ff; Gladney 2007, ss. 237ff; Giaretta 2011, ss. 200ff). Som man kanske redan nu kan ana har dessa tre strategier för- och nackdelar och skapar olika förutsättningar för det digitala bevarandet. Vilka dessa är kommer att behandlas utförligare under uppsatsens resultat-, analys- och diskussionskapitel. En nackdel med migreringsstrategin som redan nu förtjänar att nämnas då den har bäring på uppsatsens forskningsproblem, är att det i konverteringen från ett filformat till ett annat riskerar att ske förändringar i hur den fil som konverteras kan uppfattas av en användare. Exempelvis kan kursiveringar eller färger försvinna från dokument vilket potentiellt kan ändra deras informationsinnehåll. Samtidigt kan en migrering vara nödvändig för att filen ska kunna bevaras eller användas överhuvudtaget.

Digitalt arkivmaterial måste alltså kontinuerligt hanteras och beskrivas för att förbli användbart i framtiden. Inom arkivvetenskapen har teoretiska modeller tagits fram för att visa och beskriva hur förutsättningarna för ett dokument eller arkivhandling förändras under dess liv, från dess skapelse på exempelvis en myndighet till slutförvaringen i ett arkiv. De mest använda av dessa är *three ages life cycle*- och records continuum-modellerna som kopplar olika aspekter av den kontinuerliga

arkivvården till beskrivningen av dokumentens olika faser som aktiva, semi-aktiva och inaktiva (Shepherd & Yeo 2003, ss. 5ff, 8–10; Sahlén 2005, ss. 22–23). Tidigare undersökningar av autenticitet inom en digital arkivkontext har använt denna fasuppdelning, genom att identifiera den aktiva fasen som inom de system som skapade dokumenten, semi-aktiva som mellanarkiv och inaktiva som slutarkiv, med goda resultat (InterPARES 2002b).

Ofrivillig gallring och informationsförlust kan lätt uppstå både när det gäller hårdvaruhantering och metadata kring system och information som finns lagrad. Enligt RA-FS[6] 2009:1 2 kap. 1 § räknas information som gallrad vid dessa tillfällen:
"förlust av information, förlust av möjliga sammanställningar, förlust av sökmöjligheter eller förlust av möjlighet att fastställa informationens autenticitet."
Den sista punkten visar hur stor vikt som läggs vid just autenticitet inom digitalt bevarande. Autenticitet blir således något som måste kunna bevisas och aktivt dokumenteras, i en digital kontext kan den inte tas för given. Viktigt att poängtera är att autenticitet hos digitalt arkivmaterial är kopplat till huruvida förändringar uppstått i informationsinnehållet under bevarandet, men inte att informationsinnehållet är korrekt (McNeil 1998, s. 1; Duranti 2010, s. 80). Även de övriga punkterna hämtade från RA-FS 2009:1 relaterar därför till autenticitet i digitalt bevarande.

2.3 Autenticitet och det digitala samhället

Vad har då arkivens roll i samhället, allmänna handlingar och digitalt bevarande att göra med autenticiteten hos arkivhandlingar? Rättsäkerheten i samhället grundar sig på att de allmänna handlingarna är korrekt upprättade, hanterade, bevarade och att de genom detta *är vad de utges för att vara.* Detta är ett återkommande tema i flera definitioner av autenticitet (Bearman & Trent i Nilsson 2008, s. 29; Duranti 1995, s. 7; The Consultative Committee for Space Data Systems, CCSDS 2012, s. 1-9), vilket kommer att redogöras i detalj i autenticitetsdefinitionsavsnitt i uppsatsen. Som man kan ana efter den korta introduktionen till digitalt bevarande i avsnittet ovan kan problem uppstå kring att visa att innehållet i ett digitalt objekt, dokument eller allmän handling bevarats på ett sätt som inte medfört ändringar i det ursprungliga informationsinnehållet (Nilsson 2008). Vare sig om detta skett genom avsiktlig manipulation, så som i inledningens förfalskningsexempel, eller genom oavsiktligt eller oundvikligt informationsbortfall i samband med formatmigreringar eller dylikt. Autenticiteten hos objekt, dokument eller allmänna handlingar i digital form är beroende av metadata associerad till dessa i högre grad än hos analoga dokument (Lynch 2000, s. 44), även om vissa autenticitetsmarkörer[7] är gemensamma för alla arkivhandlingar, exempelvis proveniens. Dessa och andra begrepp och sammanhang relaterade till digital autenticitet kommer att undersökas ingående i uppsatsens olika delar.

---

[6] Riksarkivets författningssamling.

[7] Begreppet "autenticitetsmarkör" används i uppsatsen för att benämna alla de egenskaper eller metadata hos ett digitalt objekt som kan påvisa att objektet är autentiskt. Begreppet har tidigare använts av Cullen (2010, ss. 5ff), om än i en snävare betydelse då han endast låter det beteckna en digital stämpel som beskriver proveniens.

3. Forskningsproblem

De dokument som skapas och bevaras digitalt får aldrig en fysisk struktur. Samtidigt är det fortfarande lika viktigt som tidigare för allmänheten, regeringar, medborgare, domstolar och företag – för att inte nämna framtidens forskare – att kunna lita på de digitala dokumenten.

I bakgrunden nämndes fantasieggande siffror om den exponentiella tillväxten av digital information i världen i stort. Sverige utgör inget undantag och ser en stor tillväxt av mängden information som skapas digitalt inom offentliga förvaltningar. Riksarkivet beräknade 2011 att de 15 Pbyte[8] data som då lagrades hos statliga myndigheter skulle ha ökat till 45 Pbyte till 2015 (Riksarkivet, s. 18). En del av denna datamängd består av allmänna handlingar vilka myndigheterna enligt tryckfrihetsförordningen (1949:105) och arkivlagen (1990:782) är skyldiga att vårda, ordna och förteckna så att rätten att ta del av allmänna handlingar tillgodoses. För att handlingar ska kunna fungera som bevis på de transaktioner i vilka de uppkommit måste man kunna förlita sig på att handlingarna är äkta och inte förfalskningar eller förändrade på något oavsiktligt sätt. Autenticiteten måste säkerställas så långt möjligt. Ironiskt nog innebär informationsåldern att informationen blivit alltmer flyktig. Arkiven måste arbeta aktivt för att underhålla det elektroniska materialet på ett annat sätt än vad de måste göra med analogt material eftersom den tekniska utvecklingen och lagringsmediernas förgänglighet skapar ogynnsamma förutsättningar för bevarandet. Att det finns svårigheter i att upprätta strategier, system och rutiner för det digitala bevarandet av allmänna handlingar inom offentlig sektor i Sverige visades[9] av Riksarkivet 2010 i en enkätundersökning som distribuerades till samtliga statliga myndigheter[10] (Jarborn & Gäfvert 2010). Om detta är ett problem för myndigheter vars arkivhållning endast utgör en stödfunktion, är det intressant att veta hur problemet angrips på slutarkiven som har bevarandet som huvudsyssla. Då allt digitalt material till sist hamnar hos dessa institutioner är det en högst relevant fråga hur dessa ser på en för oss avlägsen framtid, dess användare och varför dessa människor ska kunna lita på det digitala historiska källmaterialet.

---

[8] Prefixet P, peta står för 10^{15}. För att jämföra med den globala digitala informationsmängden som nämndes i bakgrunden, 1.5 exabyte, så går det tusen petabyte på ett exabyte.

[9] Exempelvis uppgav 64 % av myndigheterna att de saknade "övergripande, dokumenterade, strategier för bevarande av e-handlingar", 26 % att de hade dessa till viss del, 5% att de inte visste och endast 5 % att de hade det (Jarborn & Gäfvert 2010, s. 8). Ett annat exempel är att på frågan "Dokumenterar ni åtgärder vid bevarande, t. ex. kontrolläsning, konverteringstidpunkt och annan vård, m.m. i enlighet med RA-FS 2009:1, 5 kap. 5§" svarade 7 % "ja", 20 % "till viss del", 20 % "känner inte till" och 53 % "nej" (Ibid., s. 16).

[10] Riksarkivets enkätundersökning bestod av 344 enkätsvar från statliga myndigheter av det ursprungliga utskicket på 387 enkäter.

4. Syfte, frågeställningar och avgränsningar

Kapitlet börjar med att undersökningens syfte och betydelse identifieras. Efter detta formuleras den forskningsfråga som används för att uppnå undersökningens syfte, tillsammans med de delfrågeställningar som måste undersökas för att besvara forskningsfrågan. På detta följer en beskrivning av de avgränsningar som gjorts samt definitioner av nyckelbegreppet autenticitet. Kapitlet avslutas med att koncepten autenticitet och tillit relateras till varandra.

4.1 Syfte

Undersökningens syfte är att utröna huruvida det finns förutsättningar för att bevara autentiska digitala allmänna handlingar inom offentliga arkivinstitutioner för tillfället och hur förutsättningarna för detta autenticitetsbevarande ser ut över ett längre tidsperspektiv. Detta görs genom att identifiera och systematisera vissa egenskaper och faktorer hos arkivhandlingarna, så kallade autenticitetsmarkörer, vilka behövs för att garantera autenticiteten över tid för en framtida användare.

Om inledningen visade hur problem med autenticitet kan uppstå i analoga källmaterial, syftar denna uppsats till att gå ett steg längre för att undersöka vilka förutsättningar en användare i en för oss avlägsen framtid kan ha för att bedöma autenticiteten hos digitala arkivhandlingar.

Uppsatsen befinner sig i ett forskningsområde som är tämligen stort, men som också utforskats mer eller mindre utförligt av många forskare från olika forskningsdiscipliner, bland andra informationsvetenskap, arkivvetenskap och digitalt bevarande. Det denna undersökning bidrar med till forskningsfältet är att undersöka vilka aspekter av autenticiteten hos digitala allmänna handlingar som bevaras av offentliga arkivinstitutioner och vilka konsekvenser detta får för framtida användare. Att utreda vilka autenticitetsmarkörer och bevarandemetoder som faktiskt används har såvitt uppsatsförfattarna vet inte gjorts tidigare i Sverige eller för svenska juridiska förhållanden. Att utreda vilka följder institutionernas val angående autenticitetsbevarandet får för framtida användare verkar inte ha studerats tidigare.

4.2 Frågeställningar

Sålunda kan den övergripande forskningsfrågan formuleras: Vilka autenticitetsmarkörer för att bevara autenticitet hos digitala objekt används av offentliga arkivinstitutioner, varför har dessa metoder/metadata valts ut och vilka konsekvenser får detta urval för de framtida användarna?

För att kunna besvara forskningsfrågan måste de följande fem delfrågorna besvaras:
1. *Vilka metoder eller vilken metadata för att bevara autenticiteten hos digitala objekt har tidigare studerats inom forskningsområdet?* Frågan angrips genom en litteraturöversikt och syftar till att åskådliggöra vilka metoder eller metadata som tidigare identifierats och vilka valmöjligheter detta medför för institutionerna.

Följande delfrågor besvaras genom intervjuer:
2. *Anser informanterna att de bevarar handlingar eller information, och hur ser de på originalitet hos elektroniska dokument?* Synen på originalitet och information kontra handling påverkar bevarandemetodiken. I en analog värld är original, arkivhandling, information och autenticitet mer sammanbundna begrepp, vilket inte är fallet i en digital kontext där mediets karaktär tvingar fram en åtskillnad mellan dessa.

3. *Vilka metoder för bevarande av autenticitet används av institutionerna och vilka av dessa ser informanterna som viktigast?* Delfrågan är nära kopplad till den övergripande forskningsfrågan och besvarar dess första del.

4. *Vilka framtida brukargrupper ser informanterna till det elektroniska materialet och vilka användningsområden och förutsättningar tror de att dessa kommer att ha?* Då uppsatsens har starkt användarfokus är frågan av högsta relevans för att utröna om de undersökta institutionerna anpassar autenticitetsmarkörerna med någon särskild framtida användargrupp i åtanke.

5. *Hur ser överlämningsprocessen, med särskilt fokus på autenticitet, av digitala objekt till de offentliga arkiven ut?* Frågan ställs då överlämningen till arkivet är en känslig process. Flytten innebär att information kan förloras eller förvanskas. Samtidigt måste mycket metadata fångas vid tillfället, metadata som kan vara svår att skapa senare om något missas.

4.3 Avgränsningar

För att undersökningen som uppsatsen behandlar skall bli genomförbar görs vissa avgränsningar gällande vilka arkivinstitutioner som undersöks och även vilket material som är aktuellt för undersökningen. Dessa beskrivs nedan och diskuteras utförligare i uppsatsens diskussionskapitel.

En avgränsning som görs är att endast offentliga arkivinstitutioner undersöks. Denna avgränsning görs för att materialet dessa ska bevara, samhällets allmänna handlingar, lyder under samma övergripande lagstiftning även om ytterligare juridiska ramverk kan påverka utöver denna på olika sätt. Samtidigt gör det att även om de studerade arkivinstitutionerna kan innehålla olika typer av arkivmaterial är förutsättningarna likartade.

Ytterligare en avgränsning görs beträffande de offentliga arkivinstitutionerna så till vida att dessa endast är mellan- eller slutarkiv, det vill säga arkiv som innehåller arkivhandlingar under den semiaktiva och inaktiva fasen i deras existens.[11] Även denna avgränsning är till för att underlätta jämförbarheten mellan arkivinstitutionerna, men främst för att uppsatsen ämnar att undersöka *arkivinstitutionernas* roll för bevarandet av autenticitet. Eftersom arkiven endast är ansvariga för det material som levererats till dem (Hirtle 2000, s. 10) måste frågeställningen bli en annan om man ska studera hur de myndigheter och förvaltningar som skapat arkivmaterialet handskats med det under dess aktiva tid. Autenticitet hos allmänna handlingar under deras aktiva fas har tidigare undersökts av Utvik (2006) och det ter sig därför som om att ett fokus på mellan- och slutarkiv skapar mer ny kunskap.

Undersökningen begränsar sig till skriftliga arkivhandlingar, bilder och deras metadata, och behandlar inte rörlig film och ljudupptagningar som också ingår i institutionernas samlingar. Detta görs för att de två sistnämnda mediatyperna inte regleras i RA-FS 2009:2, vilken är den författning som föreskriver de tekniska krav som gäller för elektroniska handlingars framställande, bevarande och överlämnande.

De autenticitetsmarkörer som identifieras i undersökningen rankas inte gentemot varandra eller graderas för att ta fram underlag för att kvantitativt bedöma autenticiteten hos digitala allmänna handlingar. Autenticitetsmarkörerna används endast för att identifiera de nödvändiga beståndsdelarna i ett system som producerar autentiska digitala allmänna handlingar. Denna avgränsning har gjorts eftersom autenticitet, och även information, i denna uppsats ses som beroende av användaren och dennes syften och förförståelse.[12] Detta gör att rankning och kvantifiering av autenticitetsmarkörerna måste göras med en specifik användare i åtanke. Intentionen bakom denna undersökning är inte att undersöka autenticitet så som den uppfattas av en särskild målgrupp, utan att ta fram ett generellt system för att besvara forskningsfrågan, vilket gör att rankningar inom systemet inte är relevanta för undersökningen.

4.4 Definitionsdiskussion

Autenticitet, och de närliggande begreppen autentiskt *record*[13] och integritet, har definierats på varierande sätt av olika institutioner, forskare och projekt, vilket gör att det kan uppstå oklarheter när man jämför olika projekt och forskningsresultat (Gränström et al. 2002, s. 17). Nedan ges en översikt av definitioner av begreppen för att visa vilka aspekter som tas upp av olika aktörer. Efter definitionerna kommer en diskussion kring vilken syn på autenticitet som ligger till grund för denna uppsats.

---

[11] Ett tänkande som härleds från *three ages life cycle*- och records continuummodellerna som nämndes i bakgrundskapitlet, avsnitt 2.2.
[12] Grunden till detta ställningstagande återges och diskuteras i avsnitt 4.4, 7.2 och 7.3.
[13] Att *record* här inte översatts beror på att det inte finns något svenskt begrepp som helt täcker in det engelska begreppets alla konnotationer. Ett utförligare resonemang kring detta återfinns i avsnitt 5.1.

Definitioner av autenticitet

The degree to which a person (or system) may regard an object as what it is purported to be. The degree of Authenticity is judged on the basis of evidence.

OAIS[14] , CCSDS 2012 s. 1-9.

/.../ addresses what the resource purports to be and how it was created.

Bearman och Trent 1998 i Nilsson 2008 s. 29.

/.../ the persistence over time of the original characteristics of the record with respect to context, structure and content.

International Council on Archives, ICA 2005, i Nilsson 2008 s. 29.

/.../ is regarded as being established by assessing the integrity and the identity of the resource.

Giaretta 2007 s. 204

Its meaning is not restricted to authentication, as in verifying authorship, but is intended to include issues of integrity, completeness, correctness, validity, faithfulness to an original, meaningfulness, and suitability for an intended purpose.

Rothenberg 2000, s. 52

Definitioner av autentiska *records*

A record is reliable when it can be treated as the fact of which it is evidence. By contrast, a record is authentic when it is the document it claims to be.

Duranti 1995, s. 7

/.../ is a record that is what it purports to be and is free from tampering or corruption.

InterPARES 1 2002c, s. 2

/.../ a) det är vad det utger sig för att vara; b) det har skapats eller skickats av den person som uppges ha skapat eller skickat det; och det har skapats eller skickats vid den tidpunkt som uppges.

SS-ISO 15489-1:2001, 7.2.2 i Hänström 2007, s. 79.

/../ as those which an authorized records creator must have originated.

The Functional Requirements for Evidence in Record keeping Project University of Pittsburgh i Gränström et al. 2002 s. 9.

/../ – same for all intents and purposes
– same functionality and relationships to other informational entities
– same "look and feel"
– same content (for any definition of the term)
– same description.

Rothenberg 2000 s. 62

[14] OAIS är en förkortning av Open Archival Information System och är en modell som kan användas för att beskriva digitala arkiv. OAIS modellen ingår i uppsatsens teoretiska perspektiv och behandlas utförligt i avsnitt 6.3.

Integritet

> The archival concept of authenticity is closely related to that of object integrity. If a physical archival object is described as having integrity, it is understood to be complete and unaltered. Integrity speaks to the object's standing in relationship to its original form whereas authenticity speaks to whether or not the object is truly what it claims to be. Whereas integrity is a relative term, authenticity is generally thought of as an established fact. The two concepts are in many ways interrelated and any discussion of authenticity will inevitably include questions of integrity.
>
> *Adam 2010, s. 596.*

> When we say a digital object has "integrity," we mean that it has not been corrupted over time or in transit; in other words, that we have in hand the same set of sequences of bits that came into existence when the object was created.
>
> *Lynch 2000, s. 30.*

Definitionerna av integritet har inkluderats då, som Adam (2010, s. 596) menar, det är ofrånkomligt att diskutera autenticitet utan att även ta upp integritet. Detta blir tydligt om man studerar exempelvis Rothenbergs definition av autentiska *records* där han menar att dessa måste ha "same content (for any definition of the term)" och "same 'look and feel'", vilka kan ses som relaterade till handlingens integritet. Giaretta (2011, s. 204) ser autenticitet som en produkt av identitet, vilket kan ses som likvärdigt med det mer arkivvetenskapliga begreppet proveniens, och integritet. Samtidigt kan ICAs och OAIS[15] definitioner av autenticitetsbegreppet ses som inbegripande ett visst mått av integritet eftersom autenticitet som beroende av att objektet är vad det utger sig för att vara inbegriper att objektet inte förändrats, och därför att integriteten är intakt.

I denna uppsats används autenticitetsbegreppet i den vidare bemärkelsen så som den beskrivs av Rothenberg (2000, s. 52) och ICA (2005) eftersom autenticitet är en del av ett system där det sker en nära växelverkan med andra aspekter som ska bevaras, så som handlingens integritet och proveniens,[16] vilket ligger i linje med Giarettas (2011, s. 204) definition. Sålunda kan uppsatsens definition kort sammanfattas: *Ett digitalt objekt är autentiskt om det digitala objektets integritet och proveniens/identitet säkrats och kan bevisas för en användare.* Då det kan vara svårt att skilja dessa aspekter ifrån varandra eftersom de kan utläsas ur samma faktorer hos handlingen ligger det en poäng i att använda det vidare begreppet. Bearman och Trents (1998) definition tar upp att det är viktigt för autenticiteten hur det digitala objektet skapades, men även att det är vad det utges för att vara, vilket också OAIS (CCSDS 2012, s. 9-1), Duranti (1995, s. 7), InterPARES 1 (2002c, s. 2) och ISO:s definitioner omfattar.

---

[15] ICA (The International Council on Archives) är en internationell fristående organisation för främjandet av olika arkivfrågor. Man samarbetar med exempelvis UNESCO och Europarådet.

[16] För att förtydliga hur autenticitet, integritet och proveniens är sammanvävda med varandra kan man göra en liknelse med ett digitalt objekt och en burk hallonsylt. Integriteten kopplad till hur hallonfröna ligger i sylten i burken, medan proveniensen är kopplad till glasburken, etiketten och bäst före-datumet. Autenticiteten bedöms utifrån om det är hallonsylt i burken precis som det står på etiketten, och om fröerna i sylten ligger som de gjorde sist de blev kontrollerade. Annars finns det en risk att det är en annan burk fastän etiketten är densamma. Alla tre aspekterna är således beroende av varandra så som bäst-föredatumet är beroende av sylten, då den beskriver en egenskap hos denna.

Autenticitetsdefinitionen som ges av OAIS är också högst intressant eftersom den anger att autenticitet kan vara graderat snarare än antingen eller och att denna gradering beror av bevis. Detta tyder på ett positivistiskt synsätt där autenticitet kan kvantifieras och mätas, men även bedömas utifrån olika måttstockar vilket ligger i linje med OAIS' fokusering på målgrupper, så kallade designated communities (CCSDS 2012 ss. 1-11, 1-13), inom det digitala bevarandet vilket kommer att återkommas till senare i uppsatsen.

Inom diplomatiken, som kommer att behandlas utförligare i uppsatsens teorikapitel, beskrivs tre olika oberoende sorters autenticitet knutna till arkivhandlingar (Duranti 1989, s.17).

Historisk autenticitet som bedöms genom att handlingen sätts in i en historisk kontext där exempelvis kronologiska fakta jämförs med handlingens *innehåll.* Exempelvis "arbetade den som påstås ha utfärdat den här handlingen vid den påstådda myndigheten med att utfärda den här sortens handlingar vid den tidpunkt som handlingen anger att den utfärdats?"

Legal autenticitet bedöms genom att studera huruvida handlingens (juridiska) utformning överensstämmer med utformningen av den sortens handlingar vid den angivna tiden för utfärdande. Exempelvis "användes denna logotyp av myndigheten vid tiden för utfärdandet och är texten skriven enligt då gällande lagar och bestämmelser?"

Diplomatisk autenticitet kommer att beskrivas utförligare i senare kapitel, kortfattat kan sägas att denna behandlar aspekter av handlingens äkthet som kan utläsas ur hur handlingen är beskaffad fysiskt och arrangerad logiskt. Exempelvis kunde dateringar av sigill i analoga dokument visa huruvida en handling kunde upprättats vid den angivna tidpunkten, vilket i en digital kontext kan ses som utbytbart med digitala signaturer enligt Duranti (2010, s. 82) .

De första två typerna av autenticitet är knutna till verifiering av handlingens faktamässiga innehåll, vilket kan vara viktigt för användare av arkivmaterialet att ta ställning till, men det är inte arkivens sak att garantera detta (Hirtle 2000, s. 10). Lynch menar även han att de två förstnämnda aspekterna av autenticitet har föga bäring för arkivvetenskapen när han skriver "An authentic document may faithfully transmit complete falsehoods" (2000, s. 36). Duranti (1995, s. 7; 2010, s. 80) och McNeil (1998, s. 1) håller med och gör en distinktion mellan trovärdigheten hos ett dokument, vilket innebär att det faktamässiga innehållet är korrekt, jämfört med autenticiteten hos dokumentet som beskriver huruvida dokumentet är vad det utger sig för att vara eller om det har förändrats jämfört med hur det ursprungligen varit. Just att det är det faktamässiga innehållet och inte informationsinnehållet är viktigt att poängtera eftersom det faktamässiga innehållet härleds från de som skapar dokumenten medan informationsinnehållet kan förändras genom bevarandeåtgärder som vidtas av arkivinstitutionerna. Som tidigare nämnts kan migreringar mellan filformat förändra hur digitala objekt visuellt uppfattas och därigenom förändra vilken information de återger. För att återknyta till diplomatikens tre autenticiteter är det dock den tredje som är mest relevant ur ett arkivvetenskapligt perspektiv eftersom

den förutom sigill och signaturer även tar upp fysiska och logiska strukturer som bidragande till en handlings autenticitet. Detta väcker frågor kring hur dessa ska bevaras i digitala handlingar eftersom ett digitalt objekt kan framställas på olika sätt, både visuellt/fysiskt och logiskt, beroende på vilken mjukvara som används för att tillgängliggöra handlingen.

4.5 Autenticitet och tillit

Tillitsfrågan inom arkivvetenskapen har fått en framträdande roll i en del artiklar om autenticitet (ex. Lynch 2000; Levy 2000; Cullen 2000; Rothenberger 2000). Ämnet gränsar till filosofi, psykologi och sociala konstruktioner och är ett eget stort forskningsfält. Denna undersökning är dock mer positivistisk i sitt närmande till ämnet och fokuserar på mätbara och konkreta delar av autenticitet.

I grund och botten bygger autenticitet på tillit till att institutionen som står för det digitala bevarandet har skött detta på ett korrekt och trovärdigt sätt (Nilsson 2008, s. 6). Oavsett hur många kompletterande och väl utvalda autenticitetsmarkörer som används för att dokumentera autenticiteten hos en digital arkivhandling blir dessa meningslösa om användaren av materialet inte kan lita på att de individer som skrivit arkivbeskrivningarna exempelvis inte ljugit eller att ett dokument verkligen hör hemma i ett visst arkiv. Gladney (2007, s. 95) menar att tillit uppkommer när en individ upplever att den känner en annan individ så väl att den kan känna säkerhet att den andre kommer att utföra det uppdrag den anförtrotts på ett ansvarsfullt och korrekt sätt, utan onödiga risker. Detta kan även förlängas till fall då den andre individen representerar eller utgörs av en institution. Tillit och autenticitet växelverkar med varandra så tillvida att utan tillit till dem som bevarat ett digitalt objekt kan autenticiteten inte bevisas fullt ut för användaren eftersom denne kommer att behöva autenticerande dokumentation i en oändlig regression för att kunna lita på materialet.[17] Samtidigt är det så att om inte hanteringen av arkivhandlingarna noggrannt beskrivs och dokumenteras inom det digitala bevarandet kan det vara svårt att upprätthålla tilliten till att en institution sköter detta korrekt. Att uppsatsen behandlar autenticitet snarare än tillit beror på att tilliten kan ses som varande på en metanivå som genomsyrar forskningsområdet, men som samtidigt kan vara tämligen svårgripbar och flyktig: tillit kan inte dokumenteras eller kvantifieras på samma sätt som autenticitet genom olika markörer. Tillit ter sig mer kopplat till ett samspel främst mellan användaren och institutionen, till skillnad från autenticitet som, även om det också påverkas av de förra, främst lägger fokus på det digitala objektet. Valet att studera den mer objektiva autenticiteten snarare än den subjektiva tilliten får konsekvenser inte bara för hur forskningsproblemet formuleras utan i förlängningen även på vilken forskningsmetodik som kan användas.

[17] För att kunna lita på, exempelvis, att en arkivbeskrivning gjorts på ett korrekt sätt behövs dokumentation som visar *hur* arkivbeskrivningen gjorts, men det behövs även dokumentation som visar att dokumentationen om hur arkivbeskrivningen gjorts skapats på ett korrekt sätt. Denna iteration måste teoretiskt sett fortgå i oändliga led, eller tills användaren uppger att de fått tillräckligt stor tillit till institutionen för att upprepningen ska upphöra.

Cullen (2000, s. 4) menar att tilliten till arkivet är ingångsläget när det gäller fysiska och analoga samlingar och skapar en känsla av att dokumenten i samlingarna är autentiska tills motsatsen bevisats. För att uppnå samma nivå av tillit gällande digitala samlingar menar Cullen att en kombination av tekniska och beskrivningsmässiga autenticeringsmetoder kombinerat med tanken om ett tredjepartsintygande[18] skulle uppfylla detta (Ibid.). Cullens inställning påvisar att tilliten måste ha en konkret underbyggnad för att vara meningsfull och kan inte existera i sig själv. Denna uppsats identifierar tillitsbegreppet som nödvändigt att ha med i en diskussion om autenticitet, men riktar sitt huvudsakliga intresse mot den tidigare nämnda underbyggnaden.

[18] Se avsnitt 7.5 om tredjepartsintygandet - arkivet som neutral garant.

5. Tidigare forskning

Inom detta kapitel presenteras empirisk forskning som bedrivits inom forskningsfälten arkivvetenskap och digitalt långtidsbevarande, både gällande digital autenticitet specifikt och andra faktorer inom det digitala bevarandet som har anknytning till forskningsområdet för denna uppsats. Forskning som berör specifika autenticitetsmarkörer återkommer istället för tydlighetens skull i resultatkapitlets litteraturöversikt eftersom det är mer knutet till uppsatsens resultat än tidigare forskning i sig. Kapitlet börjar med ett avsnitt om vad som bevaras i arkiv och digitalt bevarande och relaterar detta till autenticitet. Detta första avsnitt ger således en bakgrund till och anger forskningsläget för delfrågeställningen "Anser informanterna att de bevarar handlingar eller information, och hur ser de på originalitet hos elektroniska dokument."

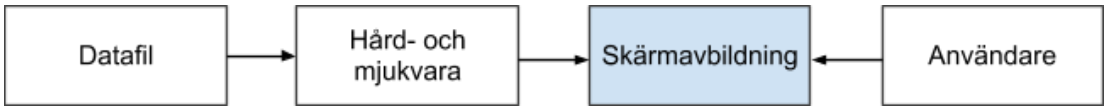
5.1 Vad bevaras i e-arkivet?

Avsnittet syftar till att med hjälp av tidigare genomförd forskning klargöra begreppen dokument, *records*, arkivhandlingar, allmänna handlingar och digitala objekt. Att klargöra olikheterna som finns mellan begreppen är viktiga för förståelsen av uppsatsens resultat, tidigare forskning, teoretiska perspektiv och dessutom för arkiv-, bevarande- och autenticitetsfrågor i stort. Jämförelsen och skillnaden mellan begreppen *records* och allmänna handlingar kommer exempelvis att ha en stor påverkan på denna undersöknings resultat.

Bucklands (1997; 1998) sammanställningar över olika synsätt på dokumentbegreppet visar att det inte finns något självklart svar för vad som utgör ett dokument. Tvärtom skiljer det sig ofta åt mellan olika vetenskapliga discipliner och tider. I översikterna kan man ana tendenser till att dokumentbegreppet blir mer inkluderande, från att ursprungligen varit bundet till det *skrivna ordet* på ett fysiskt medium, till att även inkludera exempelvis fotografier (vilka är utan text men likväl knutna till ett fysiskt medium). Detta öppnades upp ytterligare med Otlet, Schuermeyer och Briet som alla menade, möjligtvis i varierande grad, att även fysiska objekt så som arkeologiska lämningar, stenar och djur också kan vara dokument om de vidarebefordrar information till betraktaren (Otlet), utökar kunskapen hos betraktaren (Schuermeyer) eller då de tagits ur sitt naturliga sammanhang för att utgöra bevis på ett konceptuellt eller fysiskt fenomen (Briet) (Ibid.). Buckland ser dock ett behov av att öppna upp dokumentbegreppet ytterligare i digitala sammanhang eftersom digitala dokument inte kan knytas till ett särskilt medium på samma sätt som analoga dokument (1998). Han menar att det vore lämpligare att utgå från Otlets och Briets underliggande tankegångar kring dokumentens funktionalitet än deras slutgiltiga definitioner där dokumentbegreppet blir kopplat till fysiska objekt. Han menar vidare (Ibid.) att det

finns flera svar på vad som kan anses utgöra ett digitalt dokument: det kan vara återgivningen på datorskärmen eller det som återges på en pappersutskrift, eller snarare utgöras av bakomliggande funktioner[19] i datorns (hård- och/eller) mjukvara. Att det är återgivningen på datorskärmen som utgör dokumentet återkommer i performancemodellen vilken illustreras nedan (Heslop, Davis & Wilson 2002, ss. 8–10).

Figur 3. Översatt figur över performancemodellen.



Heslop et al 2002 s. 9.

Konceptet består i att användaren inte kan ta del av dokumentet eller skärmavbildningen utan att denna genomgått ett antal processer som är beroende av en datafil och hård- och mjukvara, samtidigt som alla dessa faktorer är utbytbara så länge skärmavbildningen är densamma. Alltså, det behöver varken vara "samma" datafil eller bitström[20] som en gång sparades till en viss hårdvara eller som återläses med samma mjukvara som förr, huvudsaken är att dokumentet presenteras på ett sätt likt det ursprungliga för användaren. Hur detta kan genomföras och teoribildningen kring detta kommer att redogöras för i större detalj då signifikanta egenskaper behandlas i resultatet från litteraturöversikten.

För att återknyta till dokumentbegreppet kan det konstateras att även om det används inom arkivvetenskaplig litteratur i viss utsträckning är det engelska uttrycket *records* mer beforskat inom arkivvetenskapen (Gilliland-Swetland 2005, s. 221). Det är svårt att översätta recordsbegreppet till svenska förhållanden, vilket har beskrivits och diskuterats på ett förtjänstfullt och uttömmande vis i ett flertal svenska artiklar (bland andra Hänström 2007) och böcker vilket gör att en djupare diskussion kring detta inom ramen för denna uppsats ter sig överflödig. Samtidigt är det värt att nämna några aspekter eftersom det underlättar förståelsen av den tidigare forskning och de forskningsprojekt som bedrivits. För att inte komplicera ytterligare kommer *record* inte att översättas till svenska.

En av de mest vedertagna definitionerna av recordsbegreppet är den som används inom ISO-standarden för *records* management[21] nämligen "information created, received, and maintained as evidence and information by an organization or person, in pursuance of legal obligations or in the transaction of business*". Records* är alltså

[19] Buckland (1998) exemplifierar: anta att du för 50 år sedan behövde en bok för att slå upp logaritmvärden för att göra matematiska beräkningar, men att dessa idag kan nås via internet. Det behöver inte vara självklart att dessa står utskrivna på en hemsida, utan det kan lika gärna vara en funktion i den egna datorn som beräknar dem varje gång man besöker hemsidan. Då är alltså algoritmen i sig ett dokument eftersom det genererar information, även om denna information bara genereras när hemsidan besöks.

[20] En bitström är den sekvens av ettor och nollor som utgör en enhet, exempelvis en fil, som då den processeras av en dator återger det digitala objekt som den representerar.

[21] The international standard on records management, ISO 15489

skapade och bevaras som bevis för något som har skett. Detta inneboende bevisvärde är också centralt i definitionen given av International Council on Archives (ICA) *"*A record is recorded information produced or received in the initiation, conduct or completion of an institutional or individual activity and that comprises content, context and structure sufficient to provide evidence of the activity*" (*ICA 1997, s. 22) vilken tydligt skapats med en digital kontext i åtanke. Ett *record,* oavsett analogt eller digitalt, kan därför ses som ett dokument som utgör ett bevis för att en viss händelse ägt rum. Detta medför att ett *record* som tappar sitt bevisvärde och autenticitet slutar vara ett *record,* vilket också innebär att det inte längre finns bevis för den händelse som detta *record* var sprunget ur. Var händelsen tillräckligt viktig för att dokumentera genom ett *record* bör denna bevaras på ett sätt som gör att bevisvärdet inte försvinner (Gladney 2007, s. 3), och det är därför det blir centralt med bevarandet av autenticiteten eftersom denna är tätt sammankopplad till bevisvärdet, vilket kommer att bli ännu tydligare i teorikapitlets avsnitt om diplomatik. Dessutom är det viktigt i sammanhanget att påpeka att ett digitalt *record* kan vara utspritt i flera filer i olika format i ett dokumenthanteringssystem (Duranti 2010, s. 79).

Inom svensk offentlig förvaltning motsvarar allmänna handlingar *records* till viss del.[22] Skillnaden ligger i att ett *record* måste innehålla ett bevis *för* någonting för att kunna anses vara ett *record* (inre bevisvärde), medan en allmän handling är ett bevis i sig utan att den behöver dokumentera något (yttre bevisvärde) (Hänström 2007, ss. 78–79). Att en allmän handling inte nödvändigtvis dokumenterar något beror på att enligt tryckfrihetsförordningen (1949:105 2 kap. 9 §) blir de dokument som arkiveras på en myndighet automatiskt allmänna handlingar oavsett vad de innehåller. Något som "slinker med" till arkivet kan alltså bli en allmän handling. Eftersom allmänna handlingar ska arkiveras, i de fall de inte gallras, blir dessa arkivhandlingar under den senare delen av deras levnad (Gränström, Lundqvist & Fredriksson 2000, s. 16; Hänström 2007, ss. 77–78).

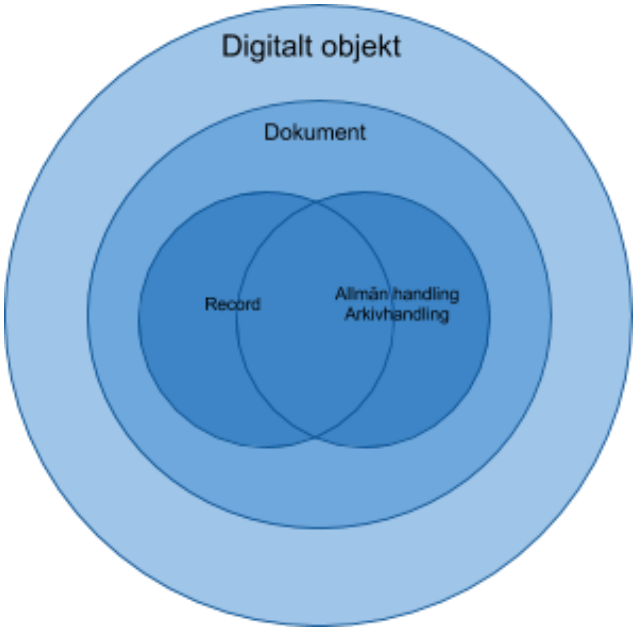
Förutom dokument, *records* och allmänna handlingar är det intressant för den här uppsatsen att tala om digitala objekt eftersom många av resonemangen berör digitalt bevarande. Därför är det viktigt att ha i åtanke hur detta begrepp relaterar till de tre förstnämnda. I de fall arkivhandlingar är digitala kommer de att bestå av en datafil, exempelvis en bildfil eller en textfil. Dessa filer är en typ av digitala objekt, det vill säga avgränsade enheter som ofta motsvaras av en bitström (CCSDS 2012, 1-1). Vinsten med att använda termen digitalt objekt snarare än digital arkivhandling är att själva metadatan som beskriver den digitala arkivhandlingen också utgör ett digitalt objekt. Detta gör att de resonemang som är generella för bevarandet av alla sorters digitala objekt[23] bättre kan beskrivas med en inkluderande term. I de fall resonemanget endast gäller digitala arkivhandlingar kommer därför denna term att användas, men eftersom de flesta resonemangen är generella för alla digitala objekt är det mer adekvat att då använda detta begrepp.

---

[22] För en fylligare analys och redogörelse kring detta i en autenticitetskontext rekommenderas Hänström 2007.

[23] Oavsett om de är enkla, som textfiler eller metadataposter, eller sammansatta som PDF-dokument eller webbsidor (Giaretta 2011, ss. 31ff).

Figur 4. Illustration över begreppsrelationerna.



För att summera: ett dokument kan vara ett *record* men är det inte nödvändigtvis, även om ett *record* alltid är ett (eller flera) dokument. Ett *record* skapat inom svensk offentlig förvaltning är alltid en allmän handling, men en allmän handling är inte alltid ett *record*. Allmänna handlingar är alltså alltid dokument.

Digitala objekt kan ses som ekvivalenta med elektroniska dokument utan tillhörande metadata. Allmänna handlingar och *records* är digitala objekt när de är elektroniska.

5.2 Bevara eller lagra – autenticitet inom digitalt bevarande

Det finns, eller har i alla fall tidigare funnits, olösta problem kring hur autenticitet ska bevaras hos digitalt arkivmaterial (InterPARES 2002b; Bradley 2006; Hänström 2007; Roeder, Eppard, Underwood & Lauriault 2008; Duranti 2010). Som tidigare visats i uppsatsavsnittet om autenticitetsdefinitionen finns ett flertal olika sätt att definiera och se på autenticitet vilket också medför att olika forskningsprojekt studerat delvis olika aspekter av området i de fall deras definitioner skilt sig åt.

Det största forskningsinitiativet inom området för digital autenticitet torde vara InterPARES, oförkortat International Research on Permanent Authentic Records in Electronic Systems, ett multinationellt forskningsprojekt som syftat till att utveckla kunskap kring långtidsbevarandet av autentiska *records* och att ta fram underlag för standarder, policys och handlingsplaner som möjliggör detta.[24] InterPARES består egentligen, till dags dato, av tre på varandra följande projekt, InterPARES 1 till 3. InterPARES 1, som pågick mellan åren 1999–2001, hade ambitionen att ta fram bland annat konceptuella ramverk, standarder och metoder för bevarande som var giltiga för vissa typer av inaktiva *records*, främst databaser och textdokument inom administrativa och juridiska organisationer (InterPARES 2002b, ss. 2–4). Projektet utmynnade bland annat i två checklistor för kriterier som ska uppfyllas för att ett *record* ska kunna anses autentiskt genom både sin aktiva och inaktiva fas (InterPARES 2000, ss. 1–9; InterPARES 2002a, ss. 1–12; InterPARES 2 2008, ss. 1–

[24] InterPARES-projektens samlade webbplats http://interpares.org/

3). Hänström undersökte 2007 huruvida dessa checklistor gick att applicera på svenska allmänna handlingar, och fann att detta inte var möjligt fullt ut (ss. 101–103). Detta på grund av att begreppen *record* och allmän handling inte helt överenstämmer i sina inre och yttre bevisvärden. Den teoretiska och konceptuella grunden bakom InterPARES 1, både checklistorna och projektet i sin helhet, låg inom arkivvetenskapen och diplomatiken (InterPARES 2002c).

Inom projektet genomfördes också fallstudier för att undersöka huruvida dessa checklistor över vad som, diplomatiskt sett, utgör ett autentiskt *record* gick att använda inom undersökningsområdets juridiska och offentliga institutioner. Man fann då att få av de undersökta dokumenthanteringssystemens innehåll kunde anses uppnå bevisgraden hos *records*,[25] och även då de gjorde detta bevarades inte dessa *records* på ett vis som även bevarade och dokumenterade deras autenticitet (Roeder, Eppard, Underwood & Lauriault 2008, ss. 121–122). För att utröna om detta var problem som var spridda inom andra samhällssektorer startades InterPARES 2, vilket pågick mellan åren 2002–2006, och som dessutom syftade till att vidareutveckla de koncept som tagits fram under det första projektet till att även gälla digitala *records* och dokument skapade inom vetenskapliga, konstnärliga och statliga sektorer (Duranti 2008, s. 1). Man fann bland annat att samma svårigheter kring bevarandet och framställandet av autentiska *records* förekom även inom dessa sektorer, samtidigt som man tyckte sig ana en förbättring och ett genomslag ute på institutionerna av de koncept som tagits fram i InterPARES 1 (Roeder, Eppard, Underwood & Lauriault 2008, s. 151).

InterPARES 3, aktivt mellan åren 2007–2012, startades eftersom man i de tidigare projekten sett tendenser till att åtgärder och strategier för att bevara autentiska *records* var situationsberoende och att en färdig mall för bevarandet blir för generell för att vara användbar om den ska vara applicerbar inom alla samhällssektorer och på alla sorters *records* eller dokument.[26] Projektet inriktade sig därför åt att skapa riktlinjer för hur små och mellanstora arkivinstitutioner bör arbeta med bevarande av autenticitet hos elektroniska *records* och dokument, eftersom dessa arkivinstitutioner har begränsade resurser för att själva ta fram riktlinjer och handlingsplaner från grunden.

De första InterPARES-projektens fynd, att det är svårt för dokument inom elektroniska dokumenthanteringssystem att uppnå recordsstatus och sedan behålla denna och sin autenticitet genom hela livscykeln, är potentiellt problematiskt eftersom det ställer frågan huruvida *records* egentligen bevaras eller endast lagras? Som tidigare nämnts i uppsatsens bakgrund är okontrollerat bortfall av information kring och i digitala objekt, vare sig det gäller deras innehållsliga struktur eller metadata som skapar en kontext eller autenticitet, en form av ofrivillig gallring (RA-FS 2009:1 2 kap. § 1). Autenticitet är ett mått på huruvida ett dokument är vad det utges för att vara och därför är *records* utan autenticitet ett ganska paradoxalt begrepp

[25] Öberg och Borglund (2006, s. 46) visade i en empirisk studie av fyra svenska myndigheter att oförmåga att nå upp till recordsbegreppets bevisvärde i elektroniska ärendehanteringssystem även förekommit i Sverige.
[26] InterPARES 3 webbplats: http://www.interpares.org/ip3/ip3_overview.cfm

eftersom *records* som inte kan bevisas vara bevis för det skeende som de är bevis för ter sig ganska meningslösa. Således kan man fråga sig om *records* eller allmänna handlingar utan autenticitet har något värde, och om de digitala objekten egentligen är bevarade eller bara lagrade om autenticiteten har fallit bort? Hänström (2007, s. 103) menar att frågor kring autenticitet hos digitala allmänna handlingar inte lyfts tillräckligt i den svenska forskningen och samhällsdebatten eftersom digitala arkivhandlingar inte begärs ut i särskilt hög grad ännu, vilket gör att deras autenticitet ännu inte börjat ifrågasättas. Enligt Hänström (s. 83) uppmärksammas autenticitetsbevarande oftast inte förrän vid överlämning till slutarkiv. Detta motsägs delvis av Utviks (2006, ss. 13ff) empiriska undersökning av svenska myndigheters sätt att tackla bevarandet av sina elektroniska allmänna handlingar där de intervjuade uppger att de arbetat mycket med just konverteringar och autenticitet. De uppgav vidare att de var medvetna om att det förelåg problem, men att vissa av dessa var svåra för dem att lösa inom myndigheterna.

Att autenticitet hos digitala dokument inte är en fråga som prioriterats av arkivinstitutioner visades av Bradley (2005) i en enkätstudie av 40 nordamerikanska kulturarvsarkiv. Studien visade tendenser till att mer objektiva autenticitetsmarkörer, särskilt teknisk metadata, inte utnyttjades i någon större omfattning utan att institutionerna istället hänvisade till sitt rykte som en källa till autenticitet (Ibid., ss. 169–170). Huruvida dessa resultat hade varit desamma om undersökningen genomförts idag, nästan 10 år senare, är osäkert.

ISO-standarden Open Archival Information System (OAIS) som beskriver funktioner och processer inom system för långtidsbevarande och e-arkiv (CCSDS 2012) har blivit så etablerad inom det digitala bevarandet att vissa projekt utgår från den för att få större relevans och giltighet. Ett multinationellt forskningsprojekt kallat SHAMAN[27] (Sustaining Heritage Access through Multivalent ArchiviNg) pågår nu med målsättningen att fördjupa OAIS-modellen och dess beskrivning av arbete med digitalt bevarande, samtidigt som man inom projektet ämnar ta fram digitala system och verktyg för bevarande (SHAMAN uå). Detta påminner om målsättningen för PLANETS-projektet[28] som avslutades 2010, där man bland annat tog fram och sammanförde verktyg för migrering till standardformat, emulering, och skapade utvärderingsverktyg som byggde på signifikanta egenskaper för om migreringar varit framgångsrika (PLANETS 2009, ss. 2–4). Även CASPAR-projektet arbetar med OAIS-modellen i åtanke. I projektets målsättningar ingår att "implementera, utvidga och validera OAIS-referensmodellen"(CASPAR 2006), vidareutveckla tekniker för att fånga information som metadata för bevarande och tillgängliggörande, undersöka och utveckla nya former för sökningar inom digitala resurser, och bidra till standardisering inom områden som berörs av CASPAR-projektet . CASPAR-projektet har även tagit fram metoder och programvara för att bedöma vissa autenticitetsaspekter hos arkivbestånd och -paket och sammanställa dessa för användare (Giaretta 2011, s. 221). Det är intressant att projekten ter sig så likartade

[27] Shamanprojektets webbplats http://shaman-ip.eu/ [2013-04-13]
[28] Preservation and Long-term Access through Networked Services, PLANETS webbplats http://www.planets-project.eu/ [2013-04-13]

när de studeras och beskrivs översiktligt, men detta kan nog vara ofrånkomligt när så många aspekter av digitalt bevarande behandlas inom samma projekt. Bland de mer specialiserade projekten kan särskilt nämnas ett par som berör områden som är intressanta för denna uppsats, nämligen InSPECT som behandlar signifikanta egenskaper,[29] vilka dessa kan vara, hur de ska beskrivas, dokumenteras och bevaras (InSPECT 2012), och InterPARES som tidigare nämnts som särskilt behandlar olika aspekter av autenticitetsbevarande.

Om man jämför målsättningarna med InterPARES 1 (InterPARES 2000, ss. 1–9) och denna uppsats kan de vid första anblicken te sig likartade. Båda syftar till att undersöka autenticitetsbevarande och respektive undersökning tar fram vissa kriterier som *records* respektive allmänna handlingar måste uppfylla för att kunna anses autentiska. Där upphör dock likheterna. Som Hänström (2007, ss. 101–103) visade kan inte checklistorna från InterPARES 1 överföras till en svensk juridisk kontext eftersom *records* och allmänna handlingar inte överensstämmer enligt de inre och yttre bevisvärdena. Det denna undersökning bidrar med i detta sammanhang är att ta fram ett system, eller checklista om man så vill, över det som måste bevaras för att *allmänna handlingar* ska vara autentiska till skillnad från InterPARES' *records*. Ytterligare en skillnad mellan InterPARES 1 och denna mastersuppsats är att fokus i denna undersökning är förskjuten till mellan- och slutarkiven och att dessa endast ska garantera det som förvaltningarna levererar till dem. InterPARES 1 delar inte detta synsätt,[30] utan är mer uteslutande inriktade på själva autenticitetsvärdet hos *records* och bekymrar sig därför inte i lika hög grad om hur detta autenticitetsvärde relaterar till den institution som bevarar dem, vilket denna undersökning gör.

[29] Ett forskningsområde som behandlas utförligare i litteraturöversiktens avsnitt 8.1.3.
[30] InterPARES-projektet verkar istället lägga över ansvaret för insamlandet av korrekt information relaterat till *records* på arkivinstitutionerna snarare än arkivbildarna när de omedelbart före sin checklista över kriterier för autenticitet hos *records* skriver "To support a presumption of authenticity the preserver must obtain evidence that:" (InterPARES 2002a, s. 5).

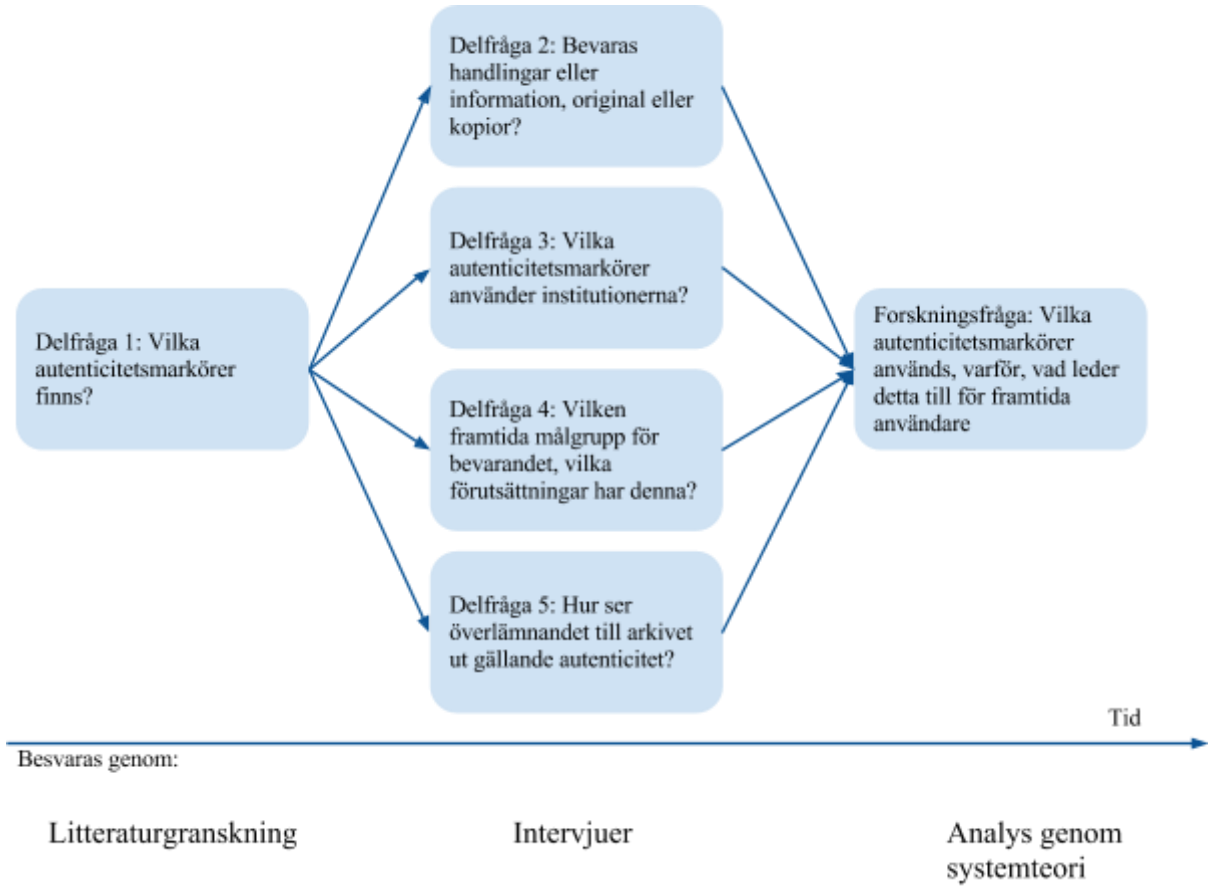
6. Metod

I detta kapitel presenteras de metoder och val som gjorts gällande forskningsmetodik och forskningsetik, och hur dessa påverkat olika aspekter av undersökningen såsom urval av institutioner, datainsamling och analys. Även om en del resonemang kring varför dessa val gjorts förekommer lämnas merparten av diskussionen kring de använda metoderna till det avslutande diskussionskapitlets metoddiskussionsavsnitt, 10.2.

6.1 Övergripande metod

Som nämndes i avsnitt 4.2 om forskningsfrågan och delfrågeställningarna bygger dessa på varandra och måste besvaras i en särskild ordning vilket illustreras i figuren nedan där respektive frågeställning kortats ned och parafraserats.

Figur 5. Illustration över frågornas relationer till metoderna.



Delfrågeställningarna besvarades med hjälp av två olika forskningsmetoder, litteraturgranskning och intervjuer. Resultaten från dessa tolkades sedan med hjälp av systemteorin, vilken beskrivs mer utförligt i kapitlet om teoretiska perspektiv, 7.1, eftersom den förutom att genomsyra forskningsmetodiken även utgör undersökningens teoretiska grundval.

6.2 Metodik gällande första delfrågeställningen

Litteraturgranskning och presentationen av tidigare forskning i vetenskapliga undersökningar syftar traditionellt till att ange forskningsfronten och att ge en översikt över den tidigare samlade kunskapen inom området (Merriam 1994, ss. 73–74). I denna undersökning används denna metod också för att besvara den första delfrågeställningen som är en förutsättning för att undersökningen ska kunna identifiera huruvida de studerade institutionerna använder alla autenticitetsmarkörer som finns tillgängliga eller om de har valt bort någon markör. Informationen och kunskapen denna svaret på denna fråga för med sig är också oumbärlig för att få god kvalitet i intervjusituationerna och intervjuresultatet som sedan företogs för att besvara övriga delfrågeställningar. Delfrågeställningen bedömdes mest lämplig att besvara genom en litteraturöversikt. Sökarbetet har varit grundligt och inspirerat av metodologin vid meta-analyser och litteraturstudier.

Två vedertagna sökmetoder för att bekanta sig med ett kunskapsläge är konsultation och manuell sökning (Backman 1998, s. 149) av vilka båda ligger till grund för litteraturöversikten. När det gäller manuell sökning gjordes till en början sökningar via databaserna LISTA och EBSCOHost och söktjänster såsom Google Scholar och SwePub efter vetenskapliga artiklar och avhandlingar för att få en överblick över det aktuella forskningsområdet. Sökningar efter monografier gjordes i Kungliga bibliotekets söktjänst Libris och efter uppsatsarbeten i Libris uppsök, en söktjänst för examensarbeten författade vid svenska universitet och högskolor. De informationsresurser denna sökning resulterade i presenteras som det första urvalet i tabellen nedan. Det andra urvalet består i de informationsresurser som tillkom efter konsultation av de källor som använts av de informationsresurser som identifierats i första urvalet. Mellan det andra och tredje urvalet sållades de informationsresurer bort som vid en närmare inspektion inte bedömdes anknyta till uppsatsämnet i sin helhet. Det tredje urvalet sammanställer hur många informationsresurser som lästes i sin helhet med målsättningen att besvara den första frågeställningen. Två gallringar gjordes sedan baserade på studiet av det tredje urvalets informationsresurser. Den första sållade bort de informationsresurser som visade sig inte anknyta till undersökningen i tillräckligt hög grad eller inte höll tillräckligt hög kvalitet, vilket resulterade i det fjärde urvalet i tabellen. Därefter sorterades de informationsresurser bort som inte kunde användas för att besvara frågeställningen, men som gav annan information som kunde användas i andra uppsatskapitel, exempelvis bakgrunden eller tidigare forskning. Övriga informationsresurser som används i uppsatsen redovisas inte här eftersom de inte ursprungligen samlats in med målsättningen att besvara forskningsfrågan.

Figur 6. Tabell över urvalsprocessen

Informations-resurs, nivå	**Vetenskaplig artikel (st)**	**Rapport (st)**	**Monografi (st)**	**Avhandling (st)**	**Masters-uppsats (st)**	**Totalt antal (st)**
Antal efter första urvalet	37	7	2	2	3	51
Antal efter andra urvalet	40	11	4	4	3	62
Antal efter tredje urvalet	35	11	4	3	2	55
Antal efter fjärde urvalet	25	9	3	3	0	40
Antal efter femte urvalet	11	6	2	2	0	21

Denna urvalsprocess påminner om den fenomenografiska analysprocess som sedan använts för att bearbeta och analysera resultaten från de fem delfrågeställningarna och som beskrivs i avsnitt 6.5.

6.3 Metodik gällande övriga delfrågeställningar

För att undersöka delfrågeställningarna 2–5 framstod ett antal olika alternativ:

Att enbart studera den dokumentation som idag skapas kring de digitala objekten på arkiven skulle kunna ge samma möjligheter att bedöma autenticiteten som den framtida användaren får. Nackdelen med detta vore att dokumentation kring framtida migreringar ännu inte skapats och att de därför inte kan utvärderas. En annan nackdel med detta förfarande vore att autenticiteten är beroende av vilket syfte användaren har. Detta sammantaget med svårigheten att hitta ett representativt digitalt arkivbestånd (eftersom samma metadataelement och samma dokumentation inte alltid nödvändigtvis kommer att uppstå) gjorde att denna metod bedömdes vara behäftad med för mycket osäkerhet för att kunna ge tydliga resultat. Den var dessutom för positivistisk eftersom den inte kan ge kvalitativa svar som varför viss dokumentation skapats, samtidigt som metanivån där tilliten för institutionen samverkar med autenticiteten blir väldigt svår att studera.

Alternativet till detta var intervjuer med personer inom arkivinstitutionerna som varit i en position att ha insyn i beslutsprocessen kring vilka bevarandemetoder och autenticitetsmarkörer som valts ut och vilken dokumentation kring dessa som skulle skapas och bevaras. Fördelen med denna metod var att undersökningen inte skedde kvantitativt på ett särskilt *reellt* bestånd utan kunde ske på ett mer teoretiskt plan där man kan undersöka vilka de ideala metadataförutsättningarna vore och även ett medelvärde kring hur det brukar vara. Nackdelen var att enligt hermeneutisk källkritik är intervjuer mindre värda som källor än vissa mer objektiva skrivna källor, i det här fallet exempelvis metadatascheman och mötesprotokoll (Alvesson & Sköldberg 1994, s. 126). I de fall det var möjligt att få tillgång till metadatascheman

och annan dokumentation var dessa ett bra underlag för att formulera mer detaljerade frågor inför intervjutillfället.

Även forskning som tidigare företagits inom undersökningsområdet pekade på att intervjuer kunde vara en bra metod. Rachel Bradley (2005) undersökte i en empirísk studie frågeställningar[31] kring autenticitet hos digitala objekt inom arkivsektorn. Hon företog dock sin undersökning genom kvantitativa enkäter, vilket medförde vissa svårigheter att använda, tolka och bearbeta materialet och resultatet. Man kunde i hennes studie ana att respondenternas syn av vad som ingick i autenticitetsbegreppet skilde sig från forskarens, och att det även fanns stora variationer inom gruppen, vilket gjorde att respondenternas svar därför inte avspeglade verkligheten eftersom de i vissa fall uppenbart misstolkade frågorna. Dessutom kunde man ana en diskrepans mellan de teoretiska resonemang som fördes inom den akademiska världen och hur hennes undersökta arkivinstitutioner arbetade. Detta fick till följd att förhållandevis många respondenter inte kunde välja några av de på förhand givna alternativen och därför fick välja "övrigt"-alternativet vilket medförde att enkätsvaren blev svårtolkade och svåra att bearbeta. Det ter sig därför som en reell risk, särskilt inom detta forskningsområde, att använda sig av metoder för att inhämta information från informanter som bygger på i förväg formulerade frågor utan möjlighet till förtydliganden eller följdfrågor, oavsett om denna metod vore renodlat kvantitativ, i form av exempelvis en enkät, eller kvalitativ, genom strukturerade intervjuer.

Eftersom möjligheten att kunna ställa följdfrågor tedde sig så avgörande för att samla in ett bra empiriskt material valdes fenomenografin som metodologisk ansats. Dahlgren och Johansson (2009, s. 122) menar att metoden passar särskilt väl för forskning som bygger på data från intervjuer av enskilda individer och som har utförts på ett halvstrukturerat sätt efter särskilda teman. Även Merriam (1994, s. 88) menar att intervjuer som inte är strukturerade är att föredra då experter ska intervjuas.[32] En intervjuteknik som förespråkas av fenomenografin, och som anammats i denna uppsats, är att intervjuarna under samtalet använt sig av så kallad *probing*, vilket innebär att ställa följdfrågor för att få fullt utvecklade svar (Ibid., s. 126). Detta var användbart i intervjusituationen eftersom det möjliggjorde att man kunde be informanterna att förtydliga vissa resonemang och begrepp, vilket gjorde att de fallgropar som upptäckts genom granskningen av Bradleys metod ovan kunde undvikas. För att vara säkra på att alla frågeteman täckts in under intervjun skapades en intervjuguide kopplad till det frågeunderlag som informanterna fick ta del av (bilaga 1). Intervjuguiden enligt fenomenografisk metod utformas efter forskarnas förförståelse och valda teorier (Ibid., s. 131), vilket passar denna undersökning utmärkt då kunskapen från litteraturöversikten, diplomatisk teori och OAIS-modellen har varit nödvändiga ingångar för att kunna ställa relevanta frågor till informanterna. Analysen enligt fenomenografisk metod presenteras i avsnitt 6.5.

---

[31] Närmare bestämt vilka metoder ett fyrtiotal enskilda arkivinstitutioner använde för att bevara autenticitet i sina digitala samlingar.

[32] Samtidigt för Eriksson-Zetterquist och Ahrne (2011, s. 40) fram en poäng i att det ofta inte är meningsfullt att göra en alltför finfördelad analys av olika intervjumetoder eftersom det ofta i intervjusamtalet kan behöva göras anpassningar beroende på informanten.

6.3.1 Institutioner och informanter – presentation och urval

Ett ändamålsinriktat urval användes för att välja arkivinstitutioner. Metoden, som även kallas kriterierelaterat urval, innebär att respondenterna väljs ut efter vissa fastställda inklusionskriterier (Merriam 1994, s. 62).

Följande inklusionskriterier för arkivinstitutioner i undersökningen har tagits fram:
1. Institutionen ska ha ett aktivt elektroniskt mellan- och/eller slutarkiv.
2. Arkivet ska vara offentligt och innehålla allmänna handlingar. I kravet ingår implicit att arkivet lyder under fastställda regler och policies av både lokal och nationell karaktär. De nationella är Riksarkivets föreskrifter samt lagtexter.

Institutionerna ska alltså vara likvärdiga gällande skyldigheter att bevara allmänna handlingar enligt arkivlagen. För att anknyta till systemteorin som kommer att presenteras i teorikapitlet, är de alltså liknande informationssystem med jämförbara förutsättningar. Informanterna har valts ut då de arbetar med e-arkiv på olika nivåer inom våra utvalda institutioner. De har delvis olika utbildningsbakgrund och ansvarsområden, men då det inte är uppsatsens syfte att jämföra utbildning och erfarenhet med deras inställning till digital autenticitet kommer dessa inte att redovisas. Vidare är informanterna pseudonymiserade och benämns informant 1–6 enligt nedan. Institutionerna beskrivs inte i detalj då det inte ingår i undersökningen att jämföra skillnader mellan dem på grund av organisation eller andra förutsättningar, utan fokus ligger på vilka metoder likartade offentliga informationssystem i Sverige väljer för att bibehålla autenticitet i digitalt långtidsbevarande.

Stockholms stadsarkiv (SSA) har uppdrag både som kommunal arkivmyndighet för Stockholms stad och som statligt landsarkiv för Stockholms län. Man har i uppbyggnaden av sitt e-arkiv en uttalad tillgänglighetsaspekt både för medborgarna och för de förvaltningar och myndigheter som behöver materialet för sina uppdrag. Arkivet är även en del av e-förvaltningen. Från SSA kommer informant 1.

System LångtidsLagring (SYLL) är ett gemensamt system för digitalt långtidsbevarande med för närvarande åtta svenska universitet och högskolor som samarbetspartners. En stor del av materialet man tar emot kommer från forskning. Universitet och högskolor har specifika och liknande krav på bevarandet av digitalt material vilket ligger till grund för samarbetet. Från SYLL kommer informant 2.

Riksarkivet (RA) har studerats både genom Riksarkivet, Landsarkivet i Lund, och Riksarkivets IT-avdelning, Stockholm, Marieberg. Landsarkiven i Sverige har tills nyligen varit självständiga, men har nu uppgått i Riksarkivets organisation. På Landsarkivet arbetar informanterna 3 och 4 och på Riksarkivets IT-avdelning informanterna 5 och 6.

6.4 Forskningsetik

Ambitionen genom hela uppsatsarbetet har varit att följa god forskningsetisk sed och ett antal hänsynstaganden relaterat till detta har därför genomsyrat arbetets gång och det slutgiltiga resultatet. En del av dessa ställningstaganden har varit sådana som kan

ses som klassiskt *forskning*setiska, det vill säga de som rör medverkande i undersökningen (även kallat extern forskningsetik), medan andra har varit de som kan ses som *forskar*etiska (intern forskningsetik) och som rör själva syftet med undersökningen eftersom det behandlar forskningshantverket och forskarens ansvar gentemot forskningen och samhället (Hermerén 2011, s. 16). Redan valet av forskningsproblem påverkades av forskaretiska ställningstaganden så tillvida att förhoppningen var att det studerade forskningsområdet skulle äga relevans för forskningen och samhället i stort.

Det går inte att komma ifrån att det föreligger ett maktförhållande i vetenskapliga undersökningar eftersom det är resultatet av forskarens val och uttolkningar som ligger till grund för den färdiga text som läsaren sedan tar del av och som informanterna presenteras i. Utav respekt för vetenskapen, informanterna och läsarna har uppsatsförfattarnas ambition varit att bedriva undersökningen så transparent och objektivt som möjligt, även om att göra detta fullt ut ur ett hermeneutiskt perspektiv är utopiskt. Under uppsatsarbetets gång har reflektioner gjorts över vilken påverkan de val som uppsatsförfattarna gjort, eller varit i färd med att göra, haft på informanter och den kontext i vilken de verkar, och även på läsaren och dennes förståelse och kunskapsbyggande. I de fall konsekvensen av valet kunnat förutses skapa obehag för någon eller på annat vis tett sig oetiskt har andra, bättre, val gjorts istället.[33] Att kunna, för sig själv eller andra, redovisa vilka val som gjorts eller kunnat göras är dock mycket svårt att genomföra fullt ut då man alltid[34] präglas av en viss förförståelse som kan göra det omöjligt att uppfatta att det faktiskt funnits valmöjligheter kring vissa aspekter.

Redan vid valet av forskningsproblem och hur detta formulerats påverkades det framtida undersökningsresultatet i och med att vissa faktorer valts ut som "problem" att studeras, medan andra faktorer ignorerats, åsidosatts eller glömts bort. Vilka val som här gjordes och hur de presenteras i bakgrunden till forskningsproblemet kan påverka hur det studerade forskningsområdet uppfattas av läsaren, något som även kan påverkas av de val som gjorts gällande vilka aspekter av tidigare forskning som berör forskningsproblemet som inkluderats och redovisats. Att institutionerna som studerats valts ut med hjälp av inklusionskriterier som varit relevanta och adekvata är en annan faktor som måste vara tydlig gentemot läsaren så att denne kan bilda sig en uppfattning om forskningens riktighet och validitet.

Ambitionen har varit att forskningsproblemet och tillhörande frågeställningar skulle vara respektfulla gentemot medverkande institutioner och informanter som delat med sig av sina kunskaper och tid. Intervjuerna har gjorts efter principen om informerat samtycke (Kvale 1997, s. 142; Svensson och Ahrne 2011, s. 30). Vid den första kontakten med samtliga informanter informerades dessa om vad undersökningen

[33] Hermerén (2011, s.18) menar att forskningsetiska avväganden ofta får göras mellan *individskyddskravet* (att individen ska slippa utstå kränkningar) och *forskningskravet* (att forskning bör bedrivas för att främja samhället och vetenskapen) och att "det inte [är] rimligt att en obetydlig skada får hindra viktig forskning*"*.

[34] Enligt ett hermeneutiskt perspektiv, vilket genomsyrar denna uppsats, något som kommer att redogöras för mer detaljerat i teorikapitlet om systemteori.

syftade till och under vilka former den och intervjuerna skulle genomföras så att informanterna var medvetna och informerade om vad de tackade ja till att medverka i. Informanterna fick även det frågeunderlag som medföljer som bilaga 1 inför intervjutillfället. För att informanterna skulle få så stor kontroll och inflytande som möjligt över sin medverkan i undersökningen spelades intervjuerna in och transkriberades, varpå transkriptionerna sändes till respektive informant så att de hade möjlighet att ändra sina utsagor. Informanterna pseudonymiserades i den färdiga texten så långt det bedömdes vara möjligt utan att deras utsagor förlorade i gravitas. Efter att det första utkastet gjorts över resultatredovisningen sändes detta till informanterna med information om vilket informantnummer de representerades av för att de skulle kunna förändra och komma med synpunkter på presentationen av deras utsagor.

Hur informanterna och de institutioner de i viss mån representerar presenterats, samt hur informanterna refererats till i texten och vilka delar av den av dem delgivna informationen som använts har också varit föremål för forskningsetiska ställningstaganden under forskningsprocessens gång. Detta eftersom hur informanterna och institutionerna porträtteras, i likhet med hur de besvarar frågorna, påverkar vilken bild läsaren får av dem (och även vilken analys som sedan kan genomföras). På grund av detta och ambitionen att leva upp till individskyddskravet var författandet av resultatredovisningen av intervjuerna tämligen besvärlig eftersom små variationer i ord kunde få stora konsekvenser. För att exemplifiera; det är stor skillnad mellan att skriva "informant X *menar"* jämfört med "informant X *resonerar*" i de fall informanter svarat på en fråga som de kanske inte var förberedda på eller kände sig helt trygga att besvara. Det har även stundtals varit svårt att formulera delar av resultatbeskrivningen eftersom förförståelsen kring frågorna och syftet med dem gjort att omedvetet lägga orden i munnen på informanterna har varit ett överhängande hot. För att undvika detta har ett ständigt reflekterande kring "vad svarar informanten" och "vad tror informanten att han eller hon svarar på" gjorts i de fall detta varit otydligt på grund av vaga formuleringar eller missförstånd från intervjuare eller informant.

Gällande resultatet hos undersökningen kan också sägas att ett undersökningsobjekt, i det här fallet en arkivinstitution, kan uppfattas i mer positiva termer om det "lever upp" till de faktorer som undersökningen i övrigt alluderat som gynsamma. Att sätta ett värderande perspektiv kring hur väl olika institutioner genomför det digitala bevarandet med fokus på autenticiteten hos de digitala objekten har inte varit en intention från uppsatsförfattarnas sida.

6.5 Resultatbearbetning och analys

De transkriberade intervjuerna och de informationsresurser som valts ut genom litteratursökningen bearbetades genom den fenomenografiska analysproceduren. Analysproceduren enligt Dahlgren och Johansson (2009, ss. 127ff) och hur den tillämpats på undersökningens två delar beskrivs i nedanstående tabell.

Figur 7. Tabell över den fenomenografiska analysprocessen.

Fas	**Funktion**	**Tillämpning på litteraturgranskningen**	**Tillämpning på intervjustudien**
Bekanta sig med materialet	Skapa förtrogenhet med materialet och se helheten.	Genomläsning av resultatet av det tredje urvalet.	Transkribering och genomläsning av intervjuerna.
Kondensation	Utskiljande av de mest signifikanta uttalandena.	Urval av citat och nyckelfraser från femte urvalets informationsresurser.	Urval av citat och nyckelfraser från intervjuerna.
Jämförelse	Likheter och skillnader i materialet gås igenom.	Uppdelning mellan bevarandestrategier och autenticitetsmarkörerna upptäcktes.	Sortering av ovanstående medförde tematisering enligt delfrågeställningarna.
Gruppering	Likheter och skillnader grupperas och relateras till varandra.	Autenticitetsmarkörerna sorterades efter funktion och teoretisk härkomst.	Vidare tematisering kring olika utsagor inom ramen för respektive delfrågeställning.
Artikulering av kategorierna	Forskaren bestämmer hur stor variationen inom en kategori kan vara utan att en ny behöver skapas.	Exempel: Migrerade och omigrerade digitala objekt som två olika kategorier identifieras.	Exempel: Tematiseringen av utsagorna inom delfrågeställningarna relaterades till det teoretiska perspektivet och svaret på delfrågeställning ett.
Namnge kategorierna	Identifiering av kategorierna.	Kategorierna definierades och rubriksattes.	Uppdelning gjordes av kategorierna enligt resonemanget ovan.
Konstrastiv fas	Forskarna kontrollerar att kategorierna är exklusiva, uttömmande och testar om olika kategorier kan slås ihop till en.	Kategoriseringen prövades och fastställdes enligt det teoretiska perspektivets uppdelning i identitet/proveniens och integritet.	Exempel: Kategorin bevarandeformat sorterades in under kategorin formatvalidering.

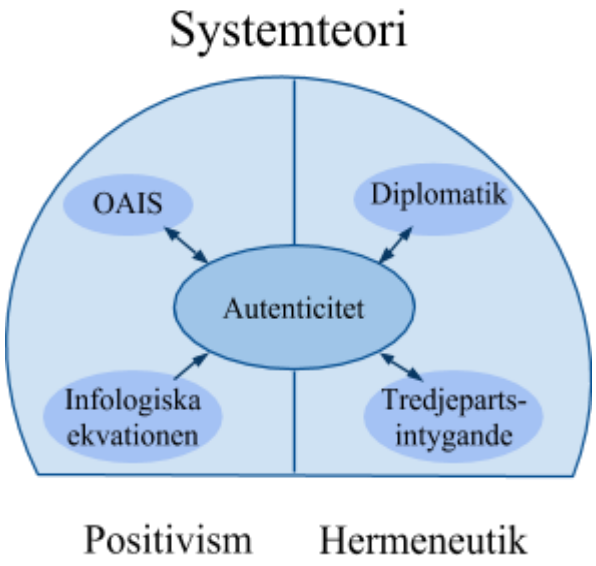
Generellt har bearbetningen av intervjumaterialet skett enligt dessa punkter. Som nämnts skapades intervjufrågorna med en relativt hög nivå av förförståelse vilket gjorde att kategorierna till stor del var givna på förhand och organiserade kring de olika autenticitetsmarkörerna.

Bearbetningsresultatet, i fenomenografin benämnt utfallsrummet (Ibid., s. 131), redovisas utan teoretisk analys i kapitlet “Resultat”. I “Analys” skrivs uppsatsens teoretiska perspektiv in i de olika kategorierna och huvudfrågeställningen besvaras med hjälp av intervjuresultaten.

7. Teoretiska perspektiv

De teoretiska perspektiv som genomsyrar denna uppsats ligger på flera nivåer. På en övergripande nivå ligger systemteorin som med dess positivistiska och hermeneutiska kombination påverkar valet av att studera autenticitet snarare än tillit, undersökningens avgränsningar, materialinsamlingen och analysen av resultatet. Inom ramen för den positivistiska halvan av det paraply systemteorin utgör ligger den infologiska ekvationen som visar vilken syn på information som används i undersökningen. Under denna halva finns också OAIS-modellen som definierar långtidsbevarande och vars metadataklassificering används för att identifiera och gruppera autenticitetsmarkörer i analysen av det empiriska materialet. Den hermeneutiska halvan av systemteorin inrymmer diplomatiken som även den används för att identifiera autenticitetsmarkörer, dock ur ett mer arkivvetenskapligt perspektiv snarare än det informations- eller systemvetenskapliga perspektiv som OAIS-modellen ger. Inom det hermeneutiska perspektivet finns även idén om tredjepartsintygandet där arkivet genom sin status som neutral part garanterar att materialet hanteras på ett korrekt sätt och att autenticiteten därmed säkras. Dessa resonemang illustreras i figuren nedan.

Figur 8. Illustration över det teoretiska perspektivet.



Alla dessa teoretiska perspektiv relaterar således till autenticitet som i denna uppsats ses som gränsöverskridande positivismen och hermeneutiken, i och med att autenticitet ses som något mätbart, men samtidigt beroende av en användare.

Kapitlet avslutas med avsnitt 7.6 där de teoretiska perspektiven binds ihop och hur de påverkar och används för att tolka de empiriska resultaten diskuteras.

7.1 Systemteori

Likt andra humanistiska vetenskaper har arkivvetenskapen oftast varit i förankrad i två inbördes motstridiga forsknings- och kunskapstraditioner. Förenklat så har positivismen sitt fokus på objekten med mätningar och förklaringar utifrån modellen orsak och verkan. Den positivistiska traditionen är nära förknippad med kvantitativa forskningsmetoder. Hermeneutiken, å andra sidan, erbjuder en mer subjektiv världsbild och sysselsätter sig gärna med fenomenologiska begrepp som olika former av mening, symboler och kommunikation. Arkivvetaren Håkan Lövblad förordade 2003 en anpassning av så kallad systemteori till arkivvetenskapen. I systemteorin finns det möjlighet för positivismens objekt och hermeneutikens fenomen att interagera (Lövblad 2003). Man kan säga att systemteori är ett pragmatiskt svar på de klassiska traditionernas begränsningar.

Teorin passar uppsatsens syfte då många begrepp som är viktiga för frågeställningarna tenderar att vävas samman vilket gör en genomgång från A till D via B och C logiskt och pedagogiskt vansklig. Kopplingarna mellan begreppen synliggörs genom att presenteras som ett system enligt systemteori. Med utgångspunkt i de teoretiska perspektiven kommer systemet in som ett sätt att sortera och beskriva verkligheten. Grundantagandet i systemtänkande är just att information om ett fenomen gradvis skapar en kontext och till slut visar ett mönster. En viktig utgångspunkt i teorin är just att delarna i systemet tillsammans, och beroende av varandra, skapar en helhet (Quisbert 2008, s. 27). Konstruktionen av systemet, vilka delar, och hur dessa arrangeras studeras. Relationer, funktioner och sammanhang, snarare än den rationalistiska världsbilden av kausalitet där krafter verkar på ting, betonas av systemtänkandet (Lundahl & Öquist 2002). Det är även av vikt att avgränsa detta system. “The main point in a systematic analysis are therefore the investigation of these functions and the demarcation of these system boundaries” (Lövblad 2003, s. 134).

I och med kopplingen till hermeneutikens tolkningstradition av ett iterativt studerande av del och helhet inom det aktuella forskningsobjektet genom den *hermeneutiska cirkeln* (Alvesson & Sköldberg 1994, ss. 116ff) får detta teoretiska perspektiv även konsekvenser på metodologin för hur tolkning och analys av resultaten av undersökningen bör utföras.[35]

Systemteori har även en koppling till funktionalistisk arkivvetenskap. Belton (2006, s. 219) menar att för att förstå det digitala dokumentet behöver man undersöka

[35] Den hermeneutiska cirkeln är ett uttryck för det grundläggande konceptet inom hermeneutiken att det inte finns någon direkt kausalitet där A påverkar B, att helheten påverkar detaljerna eller att detaljerna påverkar helheten, utan istället att A och B påverkar varandra. I denna uppsats kommer det till uttryck då vilka autenticitetsmarkörer som valts ut för att beskriva autenticiteten hos ett digitalt objekt påverkar helheten hos dettas autenticitet. Samtidigt har autenticitetsmarkörerna valts ut på grund av förutsättningarna för och synen på autenticitet hos digitala objekt.

funktionen i dess delar och hur dessa interagerar. Dessa kan gällande denna undersöknings forskningsområde vara den beskrivande metadatan, antaganden om digitala objekt och dess relationer, dataflöden etc. Vad är det i det digitala objektets natur och vilka egenskaper är det som ger ett autentiskt dokument? Frågan kan utökas till att gälla systemet som helhet – vilka karakteristika och delar i systemet kan ge autenticitet?

Vad utgör då byggstenarna i ett system? Meadows (2008, ss. 11ff) visar att ett system måste bestå av tre ting; Element (elements), sammankoppling (interconnection) och funktion[36] (function eller purpose). Som exempel kan man utgå ifrån en ubåt. Elementen kan vara sjömännen, båten, torpederna och matförrådet. Sammankopplingarna är besättningens kommunikation, de stående orderna, de fysiska lagar som bestämmer båten och besättningens rörelser etc. Elementen och sammankopplingarna resulterar i funktionen som i detta fall kan vara att verka avskräckande, försvara landet eller att upprätthålla neutraliteten.

Flexibiliteten i systemteori är stor och kan anpassas till det som kan beskrivas som ingående i systemet enligt ovanstående kriterier – element, sammankoppling och funktion. Det är framförallt förändringar i sammankoppling och funktioner som ger de mest dramatiska förändringarna i ett system, menar Meadows (2008, ss. 16ff). Besättningen kan bytas ut och det är samma system, men tar man bort funktionerna har man snarare ett museiföremål eller ett stort stycke metallskrot. Det strikta kravet på avgränsningar av systemet hindrar inte interaktion med fenomen och institutioner utanför systemet. För att en sista gång återvända till fallet med ubåten kan det gälla så skilda saker som att interagera med sjöledningscentralen, med medborgarnas förväntningar eller med bränsledepåerna. Systemteori representerar alltså en uttalat verklighetsbaserad epistemologi.

Systemets definition relaterar till uppsatsens frågeställning och kan formuleras på detta sätt:

Det offentliga e-arkivet är ett regelbaserat och formellt bevarandesystem med syftet att i ett långtidsperspektiv tillhandahålla funktionen läsbara och autentiska digitala allmänna handlingar.

Systemets funktion, sammankopplingar och element kommer att identifieras och summeras i avsnitt 7.6 där samtliga teoretiska perspektiv integreras för att klargöra vad som utgör skelettet i det ovan definierade systemet.

Quisbert (2008 s. 28) kallar dylika system för informationssystem och poängterar att systemen påverkas av krafter utifrån i form av kulturella och teknologiska förändringar. Till detta kan man även lägga politiska förändringar. Lundahl och Öquist (2002, ss. 36ff) utökar denna påverkan till att gälla åt båda hållen genom att

[36] Översättningen till svenska som "funktion" gör att begreppet blir något snävare. Litteraturen nämner både "function" och "purpose", men båda ryms inte i den svenska översättningen. Funktionen betecknar det önskade slutresultatet av systemet.

peka på det ömsesidiga beroendet mellan organisation (system) och omvärld. Som tidigare påpekats har ju systemteorin ifrågasatt det enkla kausala tänkandet där A orsakar B, eller i detta sammanhang att påtryckningar från omvärld A orsakar förändringar i system B. Meadows (2008, s. 33) benämner den ömsesidiga påverkan som en "feedback loop", med vilket hon menar att istället för att bara se hur A orsakar B måste man även ta med i beräkningen hur B också påverkar A och kanske får A att förstärka eller dra tillbaka sin påverkan.

Några svagheter med teorin, som även kan vara aktuella för denna undersökning, påpekas av Hessler (2003, s. 28) som menar att den inte helt kan följa verklighetens komplexitet och dynamik och att en tendens till vaghet blir en fara. Vidare kan det bli så att man bortser från fenomen eller variabler som kan påverka systemet om dessa faller utanför ens intresse- eller kompetensområde. Å andra sidan ligger det en poäng i förenklingen, och det som Hessler menar är vaghet öppnar också en möjlighet till tvärvetenskaplighet, vilket i denna undersökning är en styrka. Uppsatsens teoretiska perspektiv utgår alltså sammanfattningsvis från en system-funktionalistisk grundsyn. Denna är öppen och kan ta hjälp av andra vetenskaper, i detta fall från arkivvetenskap, digitalt bevarande och diplomatik. Den flera hundra år gamla diplomatiska vetenskapen behandlar frågan om autenticitet i olika former av dokument. Då den diplomatiska metoden innebär att syna dokuments egenskaper i sina beståndsdelar och analysera dess funktioner, för att sedan se vilken helhet dessa skapar, menar Lövblad (2003, s. 155) att diplomatiken är ett tidigt exempel på systemtänkande.

7.2 Den infologiska ekvationen

Inom ett informationssystem är det lämpligt att se elementet, alltså det som förädlas till resultatet eller funktionen av systemet, som information eller i vissa fall data. Den infologiska ekvationen formuleras I=i(D,S,t) där I betecknar den information som kan erhållas genom tolkningsprocessen i som är beroende av variablerna datan D, förförståelsen S och under tiden t (Langefors 1993, ss. 30ff). Detta innebär att den tillgängliga informationen inte kommer att vara densamma om förförståelsen ändras, alltså kommer två användargrupper med olika förförståelse inte att få samma information från identiska data, än mindre från data som inte är identisk. Att se användaren som en faktor eller aktör som påverkar information och dess bevarande är ett centralt koncept för denna undersökning och det är på grund av detta som den infologiska ekvationen inkluderas i det teoretiska perspektivet. Den infologiska ekvationen har tidigare använts av forskare vid LDB-centrum[37] som en teoretisk grund för att beskriva information inom en digital långtidsbevarandekontext (Runardotter, Quisbert, Jönsson, Hägerfors & Mirijamsdotter 2006, ss. 18–19) på ett mycket förtjänstfullt vis.

[37] Långsiktigt Digitalt Bevarande-centrum startade som ett forskningsprojekt i digitalt bevarande för svenska arkivinstitutioner 2004 och innefattar idag fyra parter: Riksarkivet, Kungliga biblioteket, Luleå tekniska universitet och Bodens kommun (LDB-centrum 2013).

Att användaren har en påverkan på det digitala bevarandet återkommer i definitionerna av långtidsbevarande och andra nyckelbegrepp inom OAIS-ramverket som presenteras i nästa avsnitt.

7.3 Open Archival Information System – OAIS

OAIS-modellen togs fram av flera nationella och internationella rymdforskningsinstitutioner, bland andra NASA, för att ett teoretiskt ramverk behövdes för att tillgodose behovet att ta hand om de stora mängder digitalt material som skapats i samband med rymdforskning, och modellen har sedan blivit antagen som en ISO-standard. Modellen innehåller en väl utvecklad begreppsapparat som lämpar sig att använda inom uppsatsens forskningsområde. I OAIS-modellen definieras följande begrepp:

> *Lång tid (Long term):* A period of time long enough for there to be concern about the impacts of changing technologies, including support for new media and data formats, and of a changing Designated Community, on the information being held in an OAIS. This period extends into the indefinite future.
>
> *CCSDS 2012, s. 1-12*

> *Autenticitet (authenticity):* The degree to which a person (or system) may regard an object as what it is purported to be. The degree of Authenticity is judged on the basis of evidence.
>
> *Ibid., s. 1-9*

> *Målgrupp (designated community):* An identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities. A Designated Community is defined by the archive and this definition may change over time.
>
> *Ibid., s. 1-11*

> *Oberoende förståelig (independently understandable):* A characteristic of information that is sufficiently complete to allow it to be interpreted, understood and used by the Designated Community without having to resort to special resources not widely available, including named individuals.
>
> *Ibid., s. 1-12*

Dessa fyra begrepp knyts sedan samman till OAIS-definitionen av långtidsbevarande:

> *Långtidsbevarande (Long term preservation):* The act of maintaining information, Independently Understandable by a Designated Community, and with evidence supporting its Authenticity, over the Long Term.
>
> *Ibid., s. 1-13*

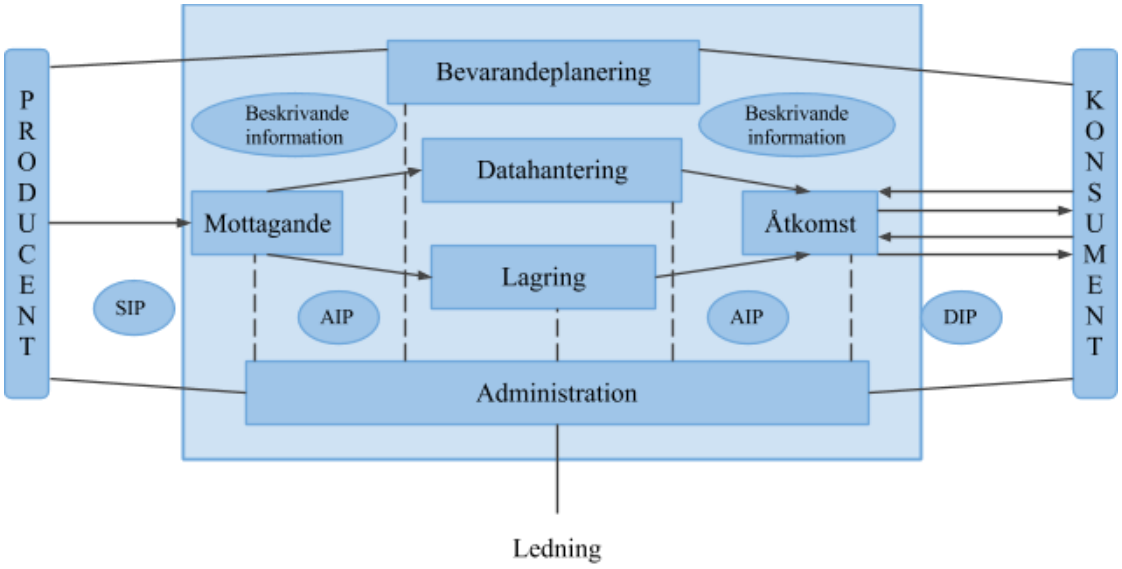
OAIS lägger alltså fokus på långtidsbevarande av *informationen*, inte handlingarna i sig. Denna information ska bevaras så att den är tolkningsbar och förståelig av en målgrupp, vilket kan ses som en spegling av hur information ses inom den infologiska ekvationen. Bevarandet är också beroende av bevis som stödjer informationens autenticitet, om detta krav inte uppfylls kan digitalt material inte anses bevarat, utan möjligen endast lagrat. Viktigt är också att informationen ska vara

oberoende förståelig, det vill säga att det ska vara möjligt att öppna, använda och förstå de digitala objekten för målgruppen.[38]

OAIS-definitionen av långtidsbevarande genomsyrar undersökningens frågeställningar och metodik på så vis att den riktar en strålkastare mot användarnas plats och behov inom långtidsbevarandet. Synen på autenticitet som beroende av bevis och ett värde hos det digitala objektet som kan graderas snarare än att vara enbart antingen eller har också starkt påverkat undersökningen. Detta bland annat eftersom det annars blir problematiskt att efter migreringar kunna bedöma ett digitalt objekt som autentiskt eftersom minsta lilla förändring uppkommen i migreringen då negerat allt autenticitetsvärde hos objektet snarare än en aspekt, exempelvis integritet.

OAIS kan användas för att beskriva ett digitalt arkiv på flera plan av vilken det mest välkända torde vara den som de flesta arkivvetare associerar till som "OAIS-modellen", nämligen den övergripande modellen över hur ett elektroniskt arkiv är uppbyggt som visas nedan.

Figur 9. Översatt figur över e-arkivets beståndsdelar.



CCSDS 2012, s. 4-1.

Denna del av OAIS-ramverket kommer inte att här redovisas i detalj då uppsatsens frågeställningar inte berör det digitala arkivets inbördes funktioner och ansvarsområden, vilket är vad denna figur visar.[39] Kort kan dock konstateras att ovanstående bild har tydliga likheter med det system som studeras enligt systemteori inom ramen för denna uppsats och att avgränsningarna mot omvärlden är desamma. Alla de element som förekommer i figuren ovan kommer att interagera och inverka

[38] Om man måste ringa en arkivmedarbetare som gått i pension men som är den ende som vet hur ett gammalt system eller maskin fungerar räknas materialet alltså inte som oberoende förståeligt bevarat.
[39] För mer information om OAIS-modellen rekommenderas Giaretta (2011, kap. 3 och 6) eller delar av Askergren (red.) 2009.

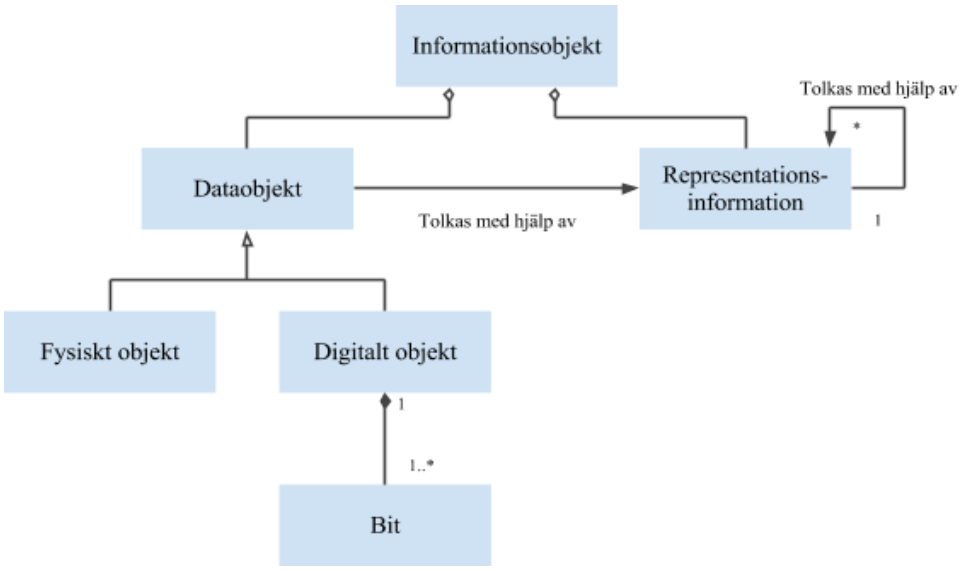
på det digitala dokumentets autenticitet då förutsättningarna för denna är beroende av aspekter som kan härledas till olika delar, eller snarare arbetsuppgifter, inom arkivet. Istället för att besvara frågeställningarna utifrån vilken arbetsprocess inom arkivet som ger upphov till respektive aspekt av autenticitet kommer uppsatsen att fokusera på *resultaten* av dessa arbetsprocesser, nämligen den metadata och de åtgärder som genereras och dokumenteras gällande handlingarna. Detta kan vara fruktbart då Gladney (2007, s. 110) menar att kvaliteten på den levererade informationen är mer intressant för användare än hur arkivet arbetat med att ta fram denna. Man skulle kunna tänka sig att om uppsatsen byggt på tillit snarare än autenticitet hade den "mänskliga faktorn" och de ingående arbetsmomenten i OAIS varit mer intressanta. Detta då man kunnat relatera till arkivets behov av att visa att de arbetar *beyond reproach* för att få tillit (Moss 2008, s. 80).

Den del av OAIS-referensmodellen som är mer intressant för denna uppsats behandlar hur metadata används för att beskriva och bevara digitala arkivhandlingar, och hur dessa kan ses som uppbyggda av olika komponenter. Detta beskrivs bäst genom att visa "nerifrån och upp" vad som i detalj bygger upp en Archival Information Package, AIP, inom OAIS-modellen.[40] Den minsta beståndsdelen inom det digitala bevarandet är en *bit* vilket motsvarar information huruvida ett värde är "av" eller "på", en nolla eller en etta (Giaretta 2011, s. 76). *En* bit kan teoretiskt sett bygga upp ett digitalt objekt, men i realiteten kan man inte få ut så mycket av en bit eftersom det krävs flera bitar bara för att beteckna en bokstav eller en siffra[41] vilket gör att en *bitström*, det vill säga många bitar efter varandra, ofta är det minsta bevarade digitala objektet.

[40] Översättningarna av begreppen som används inom denna del av OAIS-modellen har hämtats från den begreppslista (eARD 2013) som tagits fram genom delprojekt 1 i det pågående e-Arkiv och e-Diariumprojektet (eARD).

[41] Hur många bitar som behövs för att beteckna en bokstav beror på vilken teckenkodning man använder, exempelvis består 7-bitars ACSII, som namnet antyder, av sju bitar och 8-bitars ACSII åtta. Detta kan, inom det digitala långtidsbevarandet, få som effekt att ett material blir oläsbart om man inte dokumenterat vilken teckenkodning som används i ett digitalt objekt eftersom datorn då kommer att tolka ettorna och nollorna fel och inte kunna avgöra var en bokstav börjar och nästa slutar.

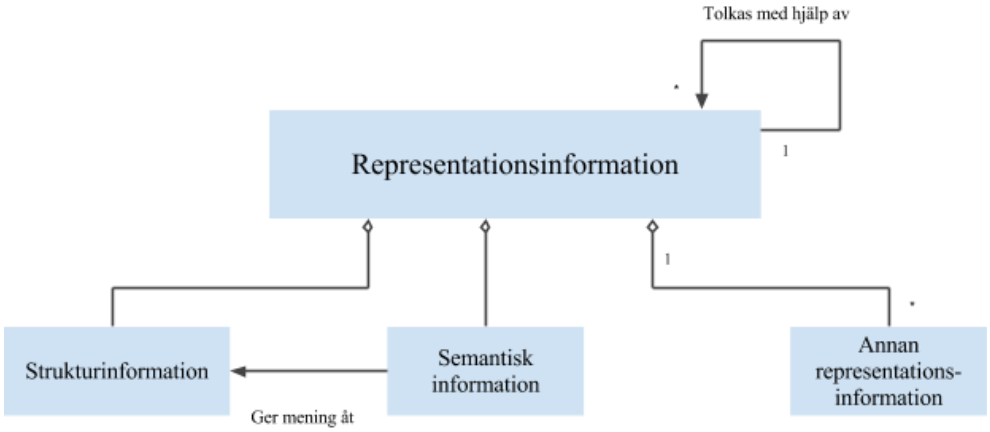
Figur 10. Översatt figur över uppbyggnaden av informationsobjekt.



CCSDS 2012, s. 4-21.

Ett dataobjekt kan bestå i antingen ett digitalt objekt eller ett fysiskt objekt, till exempel en sten eller ett papper med text på. Detta gör att OAIS-modellen i praktiken även kan användas för analoga arkiv (CCSDS 2012, s. 1-1). För att datorn ska kunna visa filen på skärmen, för att en människa ska kunna förstå vad filen innehåller eller för att arkivarien ska bli klok på varför någon har lagt en sten i arkivet måste dataobjektet tolkas med hjälp av *representationsinformation, RI* (Giaretta 2011, ss. 69ff).

Figur 11. Översatt figur över representationsinformationsobjekt.



CCSDS 2012, s. 4-23.

Representationsinformation är olika sorters metadata, som klassificeras in i tre kategorier av "information" i OAIS-modellen. *Strukturinformation* som beskriver hur dataobjektet är uppbyggt, exempelvis vilket filformat, vilken teckenkodning och hur många pixlar i höjd och bredd en bild har. Detta är främst sådan metadata som *datorn* behöver för att veta hur den ska tolka ett dataobjekt så att en korrekt representation

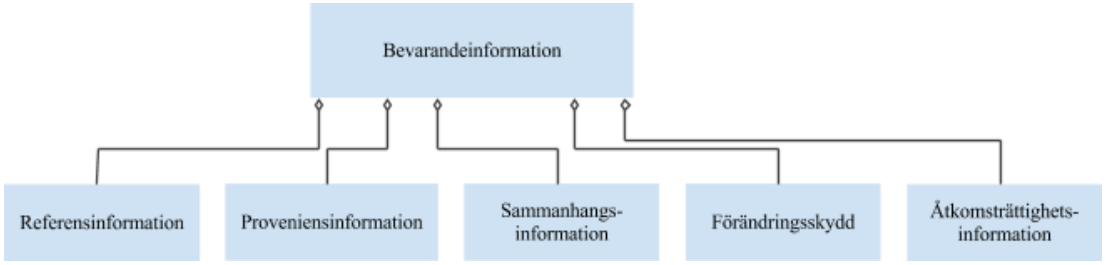
(alltså avbildning, därav namnet representationsinformation) ska visas på datorskärmen. *Semantisk information* innehåller den information *användaren* kan behöva för att förstå dataobjektet och den information om detta som finns beskriven inom strukturinformationen. Exempel är vilket språk texten i dataobjektet är skriven på, vad det innebär för filen att den är i JPEG-format, vad det är för typ av dokument och så vidare. En viktig skillnad mellan semantisk och strukturinformation är att medan strukturinformationen är beroende av IT, är den semantiska beroende av den tilltänka användaren och dess förförståelse. Detta gör att två målgrupper, designated communities, skulle behöva olika uppsättningar semantisk information medan de kan ha samma strukturinformation eftersom de teknologiska förutsättningarna för dataobjektens tolkning är desamma. Är målgruppen internationell behöver de information om att det digitala objektet är på svenska så att de vet vilken ordbok de ska använda, medan en svensk målgrupp inte ens kommer att reflektera över att de behöver veta vilket språk som används. *Annan representationsinformation* täcker in den metadata som inte lätt kan klassificeras inom de tidigare två kategorierna (Giaretta 2011, ss. 101ff). Vilken mjukvara som ska användas för att läsa ett digitalt objekt eller på vilken lagringsmedia denna är belägen, kan tjäna som exempel. Alla de tre typerna av representationsinformation kan i varierande grad påverka autenticiteten hos det dataobjekt de beskriver eftersom de har bäring på läsbarheten och därför även integriteten hos objektet.

Intressant i sammanhanget är att representationsinformation kan behövas för att förklara representationsinformationen, exempelvis att metadatan är skriven på engelska i 8-bitars ASCII. Beroende på vad man förutser sig behöva i framtiden för att göra materialet förståeligt för den tilltänkta användaren kan denna beskrivning av beskrivningen behöva upprepas i x antal led.

Då man återgår till den föregående figuren visar denne att dataobjektet tillsammans med representationsinformationen tillsammans bildar ett *informationsobjekt*, det vill säga ett tolkningsbart objekt som kan användas och förstås av en tilltänkt användargrupp. Ett informationsobjekt kan utgöras av olika sorters information, exempelvis kan ett metadataschema i sig självt vara ett informationsobjekt och inte enbart filen, eller filerna som det beskriver. Den typ av informationsobjekt som motsvarar det digitala material som e-arkivet *syftar till att bevara* kallas *innehållsinformation.* Denna beskrivs av *bevarandeinformation,* i den engelska förkortningen PDI,[42] som är sådan metadata som bör finnas för att dokumentera information som behövs för att ge upphov till tillit, tillgång och kontext för innehållsinformationen framåt i tiden (CCSDS 2012 s. 4-29).

[42] På engelska preservation description information.

Figur 12. Översatt figur över bevarandeinformation.



CCSDS 2012, s. 4-38.

De fem underavdelningar som PDI delas in i inom OAIS-modellen är:

- **Referensinformation**: den information som namnger och klassificerar innehållsinformationen inom (och utanför) arkivet. Exempelvis löpnummer/diarienummer för en allmän handling, vilken aktivitet i ett verksamhetsbaserat arkivredovisningssystem en handling är knuten till och även hela processkartläggningen i detta system så att man kan se var i systemet aktiviteten befinner sig. Mycket av denna information återkommer i de packningsbeskrivningar som medföljer ett AIP för att möjliggöra sökbarhet (Giaretta 2011, ss. 21, 178–184; CCSDS 2012, s. 4-30). Viss referensinformation behandlar autenticitet.
- **Sammanhangsinformation:** denna dokumenterar förhållandet mellan innehållsinformationen och dess omgivning och förklarar varför innehållsinformationen skapats och hur den relaterar till andra innehållsinformationsobjekt (CCSDS 2012, s. 4-30; Giaretta 2011, ss. 22, 184). Hänvisningar till andra handlingar inom ett och samma ärende, eller till lagtexter som beskriver varför en handling upprättats enligt lag kan vara exempel på sammanhangsinformation. Vissa aspekter av autenticitet täcks in av denna typ av metadata.
- **Proveniensinformation:** den metadata som beskriver innehållsinformationen från dess skapelse till uttag från arkivet; vem som skapat den, vilka ändringar som tillfogats den sedan skapelsen och vem som haft ansvar för dessa. Proveniensinformation kan ses som en subgrupp till sammanhangsinformationen, och är även mycket viktig för fastställandet av autenticiteten hos innehållsinformation (Giaretta 2011, ss. 22, 184–185; CCSDS 2012, s. 4-30).
- **Förändringsskydd[43]:** metadata som är avsedd att garantera att inga otillbörliga förändringar i innehållsinformationen skett och är övervägande tekniska till sin natur. Exempel på denna sorts metadata är checksummor, digitala signaturer och dokumentation av filformatskontroller (Giaretta 2011,

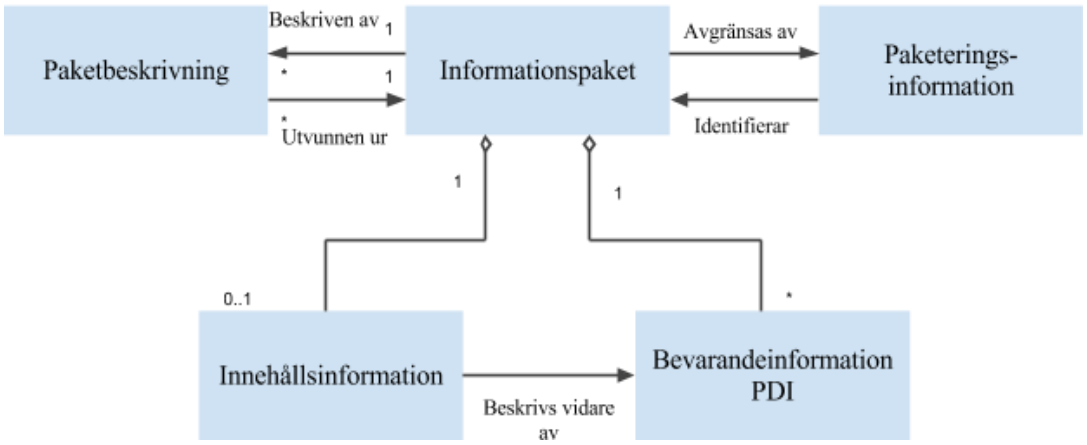
[43] I det engelska originalet kallas denna kategori för fixityinformation. Begreppet är svåröversatt eftersom fixity i det här fallet främst har att göra med fixering och på så vis oföränderlighet vilket gör att oföränderlighetsinformation därför ter sig som den mest korrekta översättningen. Delprojekt 1 inom eARD-projektet har dock översatt begreppet till förändringsskydd vilket gör att denna översättning används för att ge ökad enhetlighet i den använda terminologin.

ss. 22–23, 177–178; CCSDS 2012, s. 4-30). Är av vikt vid bedömning av autenticitet.

- **Åtkomsträttighetsinformation:** metadata om möjligheter att bruka och ta del av innehållsinformationen. Bör dokumentera sekretessbestämmelser i hänvisning till lagstiftning, och även information som rör immateriella rättigheter såsom när copyright och patentskydd förfaller (Giaretta 2011, ss. 23, 185–190; CCSDS 2012, s. 4-30). Åtkomsträttighetsinformationen är således främst knuten till användarnas och inte arkivpersonalens åtkomsträttigheter, och är därför inte av primärt intresse vid studium av autenticitet.

Även om OAIS-modellen gör en teoretisk åtskillnad mellan dessa fem metadatakategorier är det mest troligt att de kommer att förekomma huller om buller inom ett metadataschema och att deras OAIS-klassificering inte kommer att vara specificerad. I uppsatsens resultat och analys kommer de metoder arkiven använder för att bevara autentiska dokument och de metadata som genereras och dokumenteras inom dessa processer att jämföras utifrån PDI-kategorierna som ingår i OAIS-modellen för att utröna om alla autenticerande PDI-kategorier används och om vissa PDI är mer använda än andra.

Figur 13. Översatt figur över informationspaketets uppbyggnad.



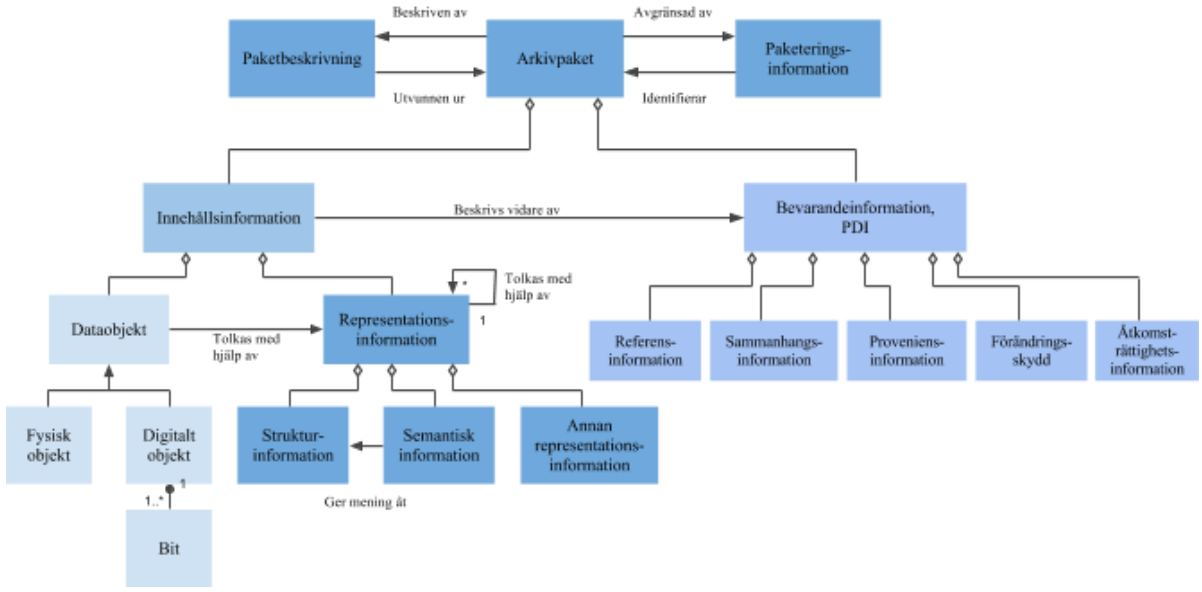
CCSDS 2012, s. 4-37.

Bilden ovan visar hur innehållsinformationen och PDI relaterar till varandra och ingår i ett *informationspaket*, det vill säga en Submission Information Package (SIP), Archival Information Package (AIP) eller Dissemination Information Package (DIP). Till varje informationspaket hör *paketeringsinformation*, som är sådan metadata som binder samman informationspaketets delar till en identifierbar enhet på någon form av lagringsmedia, till exempel att informationen ligger i en viss TAR-fil eller på en viss CD-ROM-skiva (CCSDS 2012, s. 4-32). Informationspaketet hör också samman med en *paketbeskrivning*, sådan metadata som kan liknas vid en etikett som beskriver vad som är i informationspaketet. Denna information, som ofta hämtas från PDI och innehållsinformationen, kan bestå i en arkivbeskrivning eller dylikt och är avsedd att göra innehållet i informationspaketen återsökbara i arkivet (CCSDS 2012, ss. 4-38–39). På grund av att paketbeskrivningen utvinns ur arkivpaketet fortplantar sig den

autenticitetsinformation som paketets PDI innehåller till dess beskrivning. Därför bör denna inte skilja sig från det som ingår i paketet, vilket gör att metadata från paketbeskrivningar inte kommer att studeras i detalj i analysdelen eller i systemet som helhet.[44] Hur paketbeskrivningar byggs upp beror delvis på hur själva e-arkivplattformen är uppbyggd: ett system som söker i *arkivpaketen* behöver inte lika utförliga paketbeskrivningar som ett som endast söker i dessa.

Då figurerna i avsnittet läggs samman till en enhet erhålls följande figur som beskriver hur ett arkivpaket är uppbyggt från minsta bit.

Figur 14. Översatt figur över arkivpaketet i detalj.



CCSDS 2012, s. 4-40.

Som beskrivits har representationsinformationen bäring på det bevarade objektets autenticitet i och med att det är en förutsättning för att det ska vara läsbart och att integriteten på så vis ska vara intakt. Även om förändringsskyddssinformationen från PDI-kategorin också är en garant för bevarad integritet, bland annat, kan de inte likställas fullt ut. Utan representationsinformation kan dataobjektet inte frambringas och representeras på datorskärmen, vilket gör att även om bitströmmen är bevarad så är inte informationsobjektet/innehållsinformationen det i *någon form*, autentisk eller inte. Däremot kan det vara bevarat och läsbart, och egentligen autentiskt, även om PDI:n har gått förlorad, men dess autenticitet går inte att bevisa.

Genom OAIS-metadataklassifikationen kan vissa kategorier för autenticitetsmarkörer urskiljas då dessa överensstämmer med samtliga PDI-kategorier förutom

[44] Eftersom de inte bör avvika från varandra kan intervjuer inte fånga upp huruvida de gör det eller inte, för att utreda detta måste man istället göra en kvantitativ studie där man jämför PDI och paketbeskrivningar från ett antal arkivpaket för att se om de avviker.

åtkomsträttighetsinformationen. Alltså behövs markörer för att definiera det digitala objektet eller dataobjektets referens, i form av namn eller annan identitetsmarkering, kontext som sammanhangs-PDI, oföränderlighet i form av förändringsskyddet och proveniens, i den bemärkelse att proveniensen beskriver händelser och åtgärder som vidtagits i bevarandet. De första två beskriver objektets identitet, medan förändringsskyddet beskriver integriteten hos innehållet och proveniensen är en kombination av identitet och integritet. Dessa två övergripande aspekter av autenticitet kommer nedan att behandlas och fördjupas inom en diplomatisk kontext.

7.4 Diplomatiken

Behovet av arkivteori kom ur juridiska och administrativa behov av att skapa dokument för framtida referenser. Dokumentens funktion var att utgöra ett hållbart bevis för allehanda transaktioner och ansågs få sitt bevisvärde – sin autenticitet – genom att det förvarades på vissa platser enligt fastställda administrativa procedurer (Eastwood 1994, s. 126). Ett problem med detta förfarande, som uppstod redan på kejsar Justinianus tid, var att människor började lämna in falska dokument på dessa platser – arkivet, kyrkan eller magistratet – för att på detta sätt ge de falska dokumenten en aura av autenticitet (Duranti 1998, s. 12). Under 1600-talet uppstod sålunda ett behov att uppfinna en kritisk metod och teoribildning för att råda bot på problemet med vilka samtida och medeltida dokument man kunde lita på var äkta, vilket utmynnade i diplomatiken, en av grunderna till den senare komna arkivvetenskapen.

Då diplomatiken handlar om att studera dokumentet i förhållande till dess autenticitet är det av intresse att se vilka grundelement ett *record* måste uppfylla för att vara kompletta och autenticerbara enligt diplomatisk teoribildning. Följande element måste vara klart igenkännbara (Dollar 2000, s. 22):

1. Identifiering av författare eller individ ansvarig för dokumentets skapande.
2. En handling (verb) eller manifestation av ett beslut.
3. Ett medium som gör dokumentet stabilt och därmed identifierbart.
4. Form, bestående av ingående element som datum, tiden för dokumentets skapande, själva texten och autenticering (i form av signatur).

Den andra punkten handlar om att det som undersöks ska vara en arkivhandling, en avgränsning sålunda. Detta poängteras även av Duranti (1998, s. 15), att diplomatiken ska endast studera det skrivna dokumentet, fixerat i ett medium och nedsatt med ett skrivinstrument. De tre andra punkterna visar framarbetade autenticitetsmarkörer. Duranti (2010, s. 80) delar i samband med detta upp autenticitetsresonemangen i två huvudområden – proveniens och integritet.[45] Proveniensen är nära kopplad till tanken att ett dokument är autentiskt när det är vad det utger sig för att vara. Ursprunget måste säkras och man måste ha kontroll över skapandeproceduren (Ibid.). Dollars första och fjärde punkter ovan visar detta. Integriteten å andra sidan handlar om att

[45] Duranti använder begreppen “authenticity” och “reliability”. För stringens används här begreppen proveniens och integritet. De beskriver samma sak.

dokumenten inte har blivit medvetet eller omedvetet korrumperade. Detta är beroende av hur de överförts, är de original eller kopior, bevarandesättet och en riktig ansvarskedja (Ibid., s. 81). I de följande kapitlen används Durantis uppdelning i proveniens och integritet för att visa hur dessa autenticitetsmarkörer anpassas till ett digitalt sammanhang.

7.4.1 Diplomatikens digitala anpassning

I den digitala världen finns det få forensiska eller fysiska metoder för bedömning förutom analysen av innehållet självt. Det går ju inte att analysera förhållanden som papperskvalitet, bindning, bläck och skrivstil, vilket är en självklar diplomatisk metod gällande pappersdokument. Dock finns inom diplomatiken ett system av metoder där innehållet analyseras med hjälp av vissa begrepp (Hirtle 2000, s. 23). Exempelvis anakronismer kan urskiljas i innehållet utifrån personer, social kontext, ord och uttryck och relationer, för att nämna några. Detta gränsar dock till den del av diplomatiken som är nära förbunden med historievetenskapen, men kan i vissa fall bli en fråga för arkiven. I inledningens exempel med historikern Martin Allen hjälpte arkivpersonal till att avslöja förfalskningarna genom att bland annat spåra anakronismer i form av felaktigt användande av titlar och otidsenligt användande av vissa uttryck.

Projektet InterPARES 1 utvärderade för drygt tio år sedan diplomatiken som analysverktyg inom digitalt långtidsbevarande, och menar (2002c, s. 24) att det är ett värdefullt redskap för att kunna identifiera olika implicita autenticitetsaspekter hos digitala objekt så att dessa kan göras explicita och bevaras även då de digitala objekten förflyttats ur systemet i vilka de skapades till ett arkiv. Samtidigt finns en nackdel med traditionell diplomatik då den kan vara för fokuserad på det enskilda dokumentet snarare än att se dokumentets del i arkivets kontext som viktig och meningsbärande (Storch 1998, ss. 381ff; Ibid.).

Diplomatiken har dock påbörjat en anpassning till det digitala bevarandet genom special diplomatics, eller digital diplomatics. Delmas (1996, s. 441) argumenterar för att diplomatiken ska möta den digitala utmaningen genom att helt överge det klassiska fokuset på databäraren och istället koncentrera teoribildningen om autenticitet kring informationen i dokumenten (intrinsic). Att överge databäraren skulle innebära en mindre empirisk inriktning till fördel för ett funktionalistiskt sätt att närma sig problemen. Rent praktiskt innebär det att fokus läggs på den metadata som tillhör digitala objekt och som beskriver funktioner, innehåll och identitet (Ibid).

McNeil (1998, s. 109) menar att detta synsätt förenar diplomatiken med arkivvetenskapen då diplomatiken istället för att sysselsätta sig med individuella dokument närmar sig arkivvetenskapens fokus på kontroll över arkivbestånden och arkivhandlingarnas relationer med varandra. Man närmar sig kanske lösningen på det problem som InterPARES identifierade ovan. Även om Delmas utrycker sig drastiskt när det gäller att helt släppa databäraren står det dock klart att den nya digitala diplomatiken måste handla mer om den intellektuella kontrollen av arkivmaterialet än den klassiskt fysiska.

7.4.2 Integritet

Det digitala materialets karaktär gör även att det klassiska tänkandet kring att säkra det fysiska originalet tappar något av sin mening. Diplomatikens traditionella definition av original är tämligen enkel, Duranti (1998, s.17) beskriver originalet som "Den förste i ordningen", eller som något som är "perfekt", alltså komplett och avslutat. I ett digitalt sammanhang där just överförandet och förflyttningen är det som säkrar långtidsbevarandet **(**exempelvis vid filformatsmigreringar eller då digitala objekt kopieras till nya databärare då de gamla håller på att vittra sönder) måste man dock tänka om. Då migrering innebär att man gör en reproduktion innebär detta alltid ett visst mått av förlust, vilket poängteras av Duranti (2010, s. 83). Vilken förlust som är acceptabel beror på vilken sorts dokument och information det rör sig om. Diplomatiken skiljer på intellektuella (intrinsic) och fysiska (extrinsic) egenskaper i ett dokument. Ett dokuments intellektuella autenticitet bestäms av innehållet och dess formella struktur, medan dess fysiska autenticitet kan vara egenskaper som format, färger och sigill (Duranti 1998, s. 15; Yeo 2010, ss. 90–91). Diplomatikens lösning har varit att utveckla teoribildningen kring olika sorters kopior. Att kopiering kan påverka autenticiteten, främst då integriteten hos informationsinnehållet, har gjort att man talar om tre skilda sorters kopior:

- **kopia i samma form som originalet** som innebär att två identiska kopior skapas, men att tiden mellan skapandet är det enda som skiljer dem åt. Denna typ av kopia kan uppkomma då man skriver ut två kopior av samma dokument från en dator och samma person signerar dem, de är då identiska i alla avseenden förutom att den ena skrevs ut före den andra. Denna sorts kopia är vanlig vid upprättandet av kontrakt då samma dokument med samma signaturer upprättas i ett exemplar för varje person som ingår i kontraktet.
- **imiterande kopior** är sådana som ser ut som originalet, men som kan ha förlorat några av dess karakteristika. Denna variant uppkommer exempelvis då färgen inte bevaras i kopieringen.
- **enkla kopior** uppkommer då man gör transkriptioner av dokument utan att ta hänsyn till hur de ursprungligen sett ut i sin fysiska struktur (Dollar 2000, s. 28). Detta är ett försök att släppa den klassiska diplomatiska tanken på att bevara originalet som en fysisk enhet.

Ytterligare ett viktigt koncept relaterat till integritet är ansvarskedjan, i engelskspråkig arkivvetenskaplig litteratur benämnd *chain of responsible custody*. Arkivet ska sedan de tog emot ett dokument kunna redogöra för vad som skett med dokumentet, i ett digitalt sammanhang handlar detta framförallt om att kunna beskriva hur man utfört olika sorters migreringar och åtgärder för att kontrollera att materialet inte förvanskats. Det är även av vikt att kunna se hur otillbörlig åtkomst av materialet har förhindrats och försvårats. Det gäller alltså att säkerställa att information är korrekt återgiven, inte korrekt i sig, men att den inte förvanskats, förfalskats eller bytts ut. För att få kontroll över digitala dokument när det gäller säkerhet mot både omedveten och medveten korruption föreslår Duranti (2010, s. 81) utifrån ett diplomatiskt perspektiv, att man använder standardformat, utarbetade och dokumenterade rutiner, loggar över alla förändringar och begränsad tillgång till dokumenten (access privileges)**.**

Det digitala materialets instabilitet och efemära karaktär gör att metadatan kring ansvarskedjan får stor betydelse. Thibodeau fångar elegant problemet:

> There is an inherent paradox in digital preservation. On the one hand, it aims to deliver the past into the future in an unaltered, authentic state. On the other hand, doing so inevitably requires some alteration.
>
> *Thibodeau 2002 i Adam 2010, s. 25*

Arkivhandlingarna ska alltså behållas så nära sin ursprungliga form som möjligt, och för att kunna göra detta måste man ändra på och flytta materialet. En autenticitetsmarkör blir alltså att kunna visa vad som hänt med de digitala objekten varje gång de ändrats eller flyttats. Duranti (2010, s. 84) menar att en klart bevisad och obruten ansvarsskedja är självautenticerande, och enligt diplomatiken den säkraste metoden för autenticitet i långtidsbevarande.

7.4.3 Proveniens

Problemet som diplomatiken stötte på i början av 2000-talet var just att man ur digitala objekt inte kunde härleda skapare och ursprung – proveniensen – på samma sätt som i analoga objekt (Cullen 2000, s. 3). Informationen om proveniens måste sålunda göras mer explicit i metadata som sätts till det digitala dokumentet. Proveniensen behandlar uppgifter kring dokumentets skapelse, bland annat angående arkivbildare, upphovsman eller ursprungsinstitution, i vilket sammanhang och varför dokumentet skapats. Delmas (1996, s. 442) påpekar i funktionalistisk anda att för att förstå en arkivhandlings proveniens är dokumentation kring beslutsfattande och organisation när den skapades viktigare än studiet av själva handlingen**.**

Vad detta egentligen innebär är att man inkluderar klassiska arkivteoretiska begrepp som arkivbeskrivning och archival bond i proveniensbegreppet. Archival bond innebär att relationen till andra arkivhandlingar som kommer före eller efter i en process, eller ingår i samma aktivitet, görs explicit. Detta uppnås genom arkivförteckningen och/eller processkartläggningen. Duranti (2010, s. 84) utökar resonemanget till att gälla alla medium – i hybridsystem får man alltså inte glömma att relaterade arkivhandlingar kan förekomma i både analog och digital form. McNeil (1998, s. 114) påpekar att säkrandet av proveniensen aktualiseras först i flytten till arkivet från ursprungssystemet. I ursprungssystemet är proveniensen självevident, men i ett nytt system måste denna beskrivas och säkras. Detta innebär att överföringen från ett system till ett annat är en process som måste hanteras korrekt då många möjligheter till framtida autenticering skapas här. Utvecklad metadata kring proveniens i dess här utökade betydelse är sålunda en avgörande autenticitetsmarkör.

Följande citat sammanfattar kärnfullt diplomatikens syn på autenticitet i digitalt långtidsbevarande:

> Authenticity is protected through the adoption of methods that ensure that the record is not manipulated, altered, or otherwise falsified after its creation and that it is precisely as reliable as it was when first created. It follows that an authentic electronic record is one that is transmitted in a secure way, whose state of transmission can be ascertained, that is preserved in a secure way, and whose provenance can be verified.
>
> *McNeil 1998, s. 125*

Då autenticitet säkras genom metoder som syftar till att visa att ingen förändring eller manipulering skett måste dessa metoder dokumenteras och även denna information bevaras. Denna ansvarskedja gäller både det som diplomatiken benämner proveniens och integritet, de kategorier av autenticitetsmarkörer som här kan urskiljas. Dessa överensstämmer med de kategorier som identifierades genom OAIS, även om proveniens då benämndes identitet för att undvika att blandas ihop med PDI-metadatagruppen proveniensinformation som motsvarar den dokumentation som utgörs av ansvarskedjan. Delvis länkat till ansvarskedjan och goda bevaranderutiner är tredjepartsintygandet.

7.5 Tredjepartsintygandet – institutionen som neutral garant

Arkivvetenskapen har bidragit med en annan aspekt av autenticitet – tredjepartsintygandet. Då ett arkiv kan definieras som en samling bevis ordnad som kontextuell information, inte en oorganiserad blandning, finns svaret i definitionen. Det faktum att ett dokument existerar i arkivets institutionella och sociala kontext med dess procedurer och regelverk är med denna syn ett intyg på dess autenticitet[46] (Cullen 2000, s. 3; Hirtle 2000, s. 17; Dollar 2000, ss. 25ff). Det vill säga att i och med att ett arkiv är en betrodd institution med ett gott rykte och som dokumenterat följer goda bevaranderutiner kan det stå som garant för att det kommer att bevara även detta objekt på ett pålitligt sätt. Tredjepartsintygande är således nära förknippat med tillit.

Det är dock så att olika projekt inom digitalt bevarande har valt att betona antingen metadatan eller det faktum att en betrodd institution hanterat handlingarna på ett kompetent sätt och enligt *best practice* som grundläggande för handlingarnas autenticitet (Hirtle 2000, s**s**. 18f**f**). I denna uppsats är förhoppningen att de ska kunna förenas mer harmoniskt än tidigare i och med den arkivvetenskapliga och diplomatiska grunden som behandlar tillit och dokumentets integritet kompletteras med metadatateori hämtad från OAIS-modellen.

7.6 Det teoretiska perspektivets påverkan och användning

Systemteorin har en dualism som förenar positivism och hermeneutik. Applicerat på den syn på autenticitet som genomsyrar denna uppsats får det följden att autenticitet kan studeras positivistiskt som någonting som är empiriskt mätbart, samtidigt som den hermeneutiska aspekten klargör att autenticiteten i slutändan trots allt är beroende av användaren eller betraktaren och dennes förförståelse och behov. Som tidigare nämnts är autenticitet och tillit nära förknippade med varandra. Då den hermeneutiska cirkeln appliceras på de två begreppen blir det tydligt att autenticitet är en del av tillitens helhet eftersom tillit alltid är en förutsättning för autenticitet, medan

[46] Problematiskt med tanke på exemplet med historikern Martin Allen i inledningen som smugglade in dokument med falsk proveniens bland autentiska dokument.

autenticitet endast är en av många faktorer som kan påverka tilliten. Även om en avgränsning gjorts inom undersökningen att studera autenticitet snarare än tillit innebär detta att tillit inte är något man kan, eller bör, bortse ifrån. Samtidigt är det positivistiska studiet av autenticitet det som är själva fokuset eftersom det är den enda faktorn man vid *denna tidpunkt* kan studera för att besvara forskningsfrågan. Framtidens användare och vilka möjligheter de kommer att ha för att bedöma autenticiteten hos digitala allmänna handlingar och objekt kommer ytterst att vara beroende av tilliten de har till institutionen som bevarat dessa, men detta är inget som kan mätas eller ens studeras i vår samtid. Det som kan studeras nu är trots allt endast de förutsättningar som existerat fram till denna tidpunkt och tilliten till en arkivinstitution skulle teoretiskt sett kunna vara fullgod nu – men en skandal om exempelvis fem år skulle kunna rasera den totalt. Den positivistiska autenticiteten hos skandalinstitutionens samlingar skulle dock kunna vara intakt, såvida inte skandalen vore att den positivistiska autenticiteten förlorats. Samtidigt, som en följd av del–helhetsperspektivet, studeras *vissa förutsättningar* för tillit i och med att den positivistiska autenticiteten studeras.

Användandet av systemteori påverkar både avgränsningen av forskningsområdet, forskningsmetodologin som kan användas och vidare teoretiska ramverk. Inom det paraply som systemteorin kan sägas utgöra samlas den infologiska ekvationen, OAIS-modellen, diplomatiken och tredjepartsintygandet för att tillsammans belysa olika delar i det system undersökningen ämnar att identifiera och förklara. Genom forskningsfrågan identifieras den önskvärda funktionen, det vill säga slutprodukten, som i det här fallet utgörs av autentiska digitala allmänna handlingar. Systemet som producerar denna funktion identifieras genom undersökningens avgränsningar som offentliga svenska e-mellan- och slutarkiv, vilka lyder under liknande juridiska, ekonomiska och teknologiska förutsättningar för att systemets ingående element ska kunna vara likartade varandra. Inom systemet förädlas element genom interaktion med olika sammankopplingar för att slutligen resultera i systemets funktion. Elementet i undersökningens system är *information* medan sammankopplingarna är olika *autenticitetsmarkörer*. Genom att autenticitetsmarkörerna fylls med information, liknande hur rören i ett hydrauliskt system fylls och leder vätska till sin slutdestination, samverkar informationen och autenticitetsmarkörerna till att systemets funktion, den autentiska digitala allmänna handlingen, erhålls. Viktigt att hålla i åtanke är att precis som det hydrauliska systemet inte kan uppnå sitt syfte, till exempel att ge en viss lyftkraft, om vätskenivån är för låg, kommer inte undersökningens e-arkivsystem att producera sin funktion om informationen saknas eller är bristfällig i de olika autenticitetsmarkörerna.

OAIS-modellen, både över e-arkivet som helhet och metadatan som hör till digitala objekt, utgör ett system i sig själv och är därför tacksamt att använda inom det system som identifieras inom undersökningen. OAIS ger dessutom en bra grund för den positivistiska aspekten av systemet eftersom modellen definierar autenticitet som något som är mät- och graderbart och beroende av bevis. Diplomatiken identifierar delvis samma autenticitetsmarkörer som OAIS-modellen, men sätter dessa i en mer arkivvetenskaplig kontext. På så vis vävs också in det mer hermeneutiska perspektivet där tilliten till institutionen är central, i och med tredjepartsintygandet, och där autenticitet är något som är beroende av användaren. Det senare perspektivet återkommer också i OAIS-modellens definition eftersom det delvis är olika

användares syn på autenticiteten hos ett digitalt objekt som skapar graderingen. Att diplomatiken och OAIS-modellen har liknande autenticitetsmarkörer är på intet sätt förvånande eftersom det i OAIS-modellens definition på långtidsbevarande ingår att autenticitet och bevis för denna ska bevaras och att autenticitetsavgörande och -bevarande är diplomatikens främsta mål. Då OAIS-modellen skapats i en delvis arkivvetenskaplig bevarandekontext vore det snarare anmärkningsvärt om de inte till stor del överensstämde.

De övergripande aspekterna hos det digitala objektet som måste bevaras och dokumenteras för att bevara och bevisa autenticitet enligt OAIS och diplomatiken är proveniens/identitet och integritet. Dessa är endast en övergripande kategorisering som genom litteraturgranskningen och intervjuerna fylls med faktiska autenticitetsmarkörer som används av de studerade institutionerna. Detta möjliggör i sin tur att systemets sammankopplingar kan identifieras och att systemet i sin helhet kan kartläggas. För att klargöra vilka valmöjligheter de studerade arkivinstitutionerna haft kring autenticitetsmarkörer kommer den härpå följande litteraturgranskningen att redogöra för vilka metoder som tidigare identifierats i den vetenskapliga litteraturen och hur dessa bevarar autenticitet.

8. Resultat

Resultatredovisningen är uppdelad efter delfrågeställningarna som formulerats för att besvara forskningsfrågan. Den första delfrågeställningen behandlas med hjälp av en litteraturöversikt under 8.1 med underrubriker. Delfrågeställningarna 2-5 som besvarats med intervjuer som metod finns under avsnitt 8.2-8.5. Forskningsfrågans första del "vilka autenticitetsmarkörer för att bevara autenticitet hos digitala objekt används av offentliga arkivinstitutioner" avhandlas i detta kapitel.

8.1 Hur bevaras autenticitet?

Vilka metoder eller vilken metadata för att bevara autenticiteten hos digitala objekt har tidigare studerats inom forskningsområdet?

I avsnitt 8.1.1 presenteras de tre vedertagna metoder som finns för att kunna bevara digitala objekt under mycket långa tidsrymder. Frågan om hur dessa metoder förhåller sig till autenticitet följs upp i efterföljande avsnitt. Vid bevarande av omigrerat arkivmaterial används autenticitetsmarkörer som beskrivs i 8.1.2. Efter migrering av objekten får begreppen signifikanta egenskaper och kopior stor betydelse, dessa beskrivs i 8.1.3. Gemensamt för både migrerat och omigrerat material är nödvändigheten att beskriva proveniensen, vilket tydliggörs i diplomatikavsnittet i uppsatsens teoridel.

8.1.1 Digitalt bevarande

Digitalt bevarande som forsknings- och arbetsfält inbegriper strategier, åtgärder och metoder för hur man bevarar digitala objekt, oavsett om dessa är efter analoga förlagor, det vill säga digitaliserade, eller born-digital, alltså skapade digitalt utan analoga förlagor (Digital Preservation Testbed Project 2003, ss. 8–10; Galloway 2004, ss. 553ff).

Som tidigare nämnts i bakgrundskapitlet brukar man tala om tre huvudsakliga strategier för digitalt bevarande.[47] Den första av dessa är museimetoden som består av att både det digitala objektet i dess ursprungliga bitström, den mjukvara som användes för att skapa och läsa denna och även den hårdvara som användes till detta

[47] I vissa källor, exempelvis Digital Preservation Testbed Project 2003, förekommer fler strategier, men dessa utgör snarast undergrupperingar och finfördelningar av de tre strategier som beskrivs här och är inte unika strategier i sig.

bevaras (Digital Preservation Testbed Project 2003, s. 8). Fördelen med denna metod är att det digitala objektet kan upplevas helt och hållet som vid dess tillkomst, eller under dess aktiva period så som den benämns enligt records life cycle- eller records continuummodellerna. Nackdelen är att det blir oerhört kostsamt och besvärligt för en arkivinstitution att bevara datorskärmar, möss, högtalare, hårddiskar, disketter med mera i ett brukbart skick tillgängligt för användare. Både på grund av att hårdvaran har en begränsad livslängd och kommer att gå sönder, men också för att det blir enormt många kombinationer av hård- och mjukvara som måste sparas för att alla digitala objekt som skapats av en arkivbildare (eller flera vilket ofta är fallet för offentliga arkivinstitutioner som tar emot arkivmaterial från många olika myndigheter) ska kunna användas med dess ursprungliga förutsättningar. Dessutom innebär det att det digitala arkivmaterialet endast kan återläsas och brukas då användaren är fysiskt på arkivinstitutionen. Trots detta används metoden enligt Galloway (2004, s. 557) av vissa konstmuseer för att bevara digital konst.

Den andra metoden, emulering, bevarar det digitala objektets bitström intakt och även mjukvaran som tagit fram eller ursprungligen använts för att visa objektet (Digital Preservation Testbed Project 2003, ss. 10ff; Galloway 2004, ss. 574ff; van der Hoeven, J. & van Wijngaarden, H. 2005; Gladney 2007, ss. 238ff; Giaretta 2011, ss. 197ff). Till skillnad från museimetoden bevaras dock inte någon hårdvara. Fördelen med denna metod, liksom museimetoden, är att bitströmmen är oförändrad och att den representation som återges är exakt densamma som under den aktiva fasen. Nackdelen är att det krävs att nya program ständigt tas fram för att kunna läsa den gamla mjukvaran så som den såg ut i den aktiva fasen eftersom nya operativsystem och ny hårdvara kan göra att det gamla programmet kanske inte ser likadant ut, eller i värsta fall inte längre går att köra. Dessa program kan vara kostsamma att ta fram både i programmeringstimmar och licensavgifter, och även här kan uppstå en situation där man måste ta fram många kombinationer av program och operativsystem.

Den tredje strategin, migrering, bevarar inte alltid bitströmmen hos det digitala objektet och inte heller någon relaterad mjuk-eller hårdvara, utan syftar helt till att föra över samma informationsinnehåll i det digitala objektet till nya format som är kompatibla med den nya mjuk- och hårdvara som utvecklas (vilket tillsammans med att databärarna vittrar sönder är det som skapar problemet som de tre bevarandestrategierna försöker lösa) (Galloway 2005, ss. 574ff; Gladney 2007, ss. 237ff; Giaretta 2011, ss. 200ff). Migreringar kan ske på en rad olika sätt och med varierande konsekvenser avseende integriteten hos det digitala objektets bitström, informationsinnehåll eller kontext. Att överföra elektronisk data från en databärare till en annan utan att samtidigt förändra strukturen hos det digitala objektet eller datapaketet är en typ av migrering som kallas *replikation* (vilket exempelvis sker vid överföring av information från en hårddisk till ett USB-minne). Att behålla bitströmmen hos de digitala objekten intakt samtidigt som förändringar eller tillägg görs i paketstrukturen kallas *refreshment* (kan ske då information flyttas till andra mappar eller system) (Giaretta 2011, s. 200). Då dessa typer av migreringar inte ändrar själva *datainnehållet* på något sätt utan endast byter ut lagringsmedia är sannolikheten för att information ska försvinna i migreringen tämligen låg. Risken för detta är avsevärt mycket högre vid migreringar mellan olika dataformat (exempelvis då ordbehandlingsdokument konverteras från ett word- till ett PDF-format) eftersom

formatförändringen medför att bitströmmen inte längre är intakt då filformaten medför olika regler för i vilken ordning ettorna och nollorna ska befinna sig i bitströmmen för att bilden på skärmen ska se likartad ut. Migreringar som förändrar bitströmmen hos *metadatan* kallas *ompaketering* medan migreringar som medför förändringar för det *digitala objektet* och eventuellt även metadatan kallas *transformering* (Ibid., s. 201). Transformeringar är den typ av händelser som potentiellt skapar störst problem för bevarandet av autenticitet inom digitala arkiv eftersom de innebär en reell förändring på bitströmsnivå av det material som ska bevaras. Något som i sin tur kan få följder i hur materialet kan uppfattas, tolkas och användas. Förutom att den del av metadatan som är kopplad till det gamla formatet kommer att försvinna eller förändras finns det även en risk att så som det digitala objektet ser ut på skärmen, dess så kallade representation, inte kommer att vara densamma före och efter transformationen. Detta kan ge följder för det digitala objektets informationsinnehåll om förändringar skett vilka påverkar tolkningen av datan, till exempel kan kursivering, fetstil eller färger vara väldigt meningsbärande i vissa dokument och att delar av informationen kan försvinna tillsammans med dessa. I denna uppsats ses informationsinnehållets integritet som nära besläktad och sammankopplad med autenticiteten hos det digitala objektet vilket gör att transformeringar kan påverka dessas autenticitet i de fall informationsinnehållet förändras okontrollerat.[48] För att migreringsstrategin ska kunna bevara autentiska digitala objekt har metoder tagits fram för att dokumentera och bevara olika aspekter av autenticitet, dessa redovisas i de kommande avsnitten.

8.1.2 Autenticitetsmarkörer

Ett antal olika tekniker, arbetssätt eller metoder har tagits fram för att kunna säkerställa autenticitet hos digitala dokument och objekt. Det de har gemensamt är att de skapar dokumentation kring de digitala objekten som kan sparas som metadata i relation till dessa. En teknik som tagits fram som är kopplad till integriteten är användandet av algoritmer som beräknar ett "värde"av ett digitalt objekt, dessa värden brukar kallas *checksummor*, men även hashvalues, message digests eller bara digests.[49] De digitala objekten kan i det här fallet både vara enskilda filer i olika format eller den metadata som hör till filerna. Genom att räkna ut och sätta checksummor till digitala objekt kan man garantera att deras innehåll inte förändrats (Gladney 2007, s. 165: Adam 2010, ss. 600ff: Giaretta 2011, ss. 214ff). För att kontrollera huruvida ett digitalt objekt förändrats beräknas checksumman ånyo med samma algoritm som tidigare och denna nya checksumma jämförs sedan med den första, om dessa är identiska är det digitala objektet också identiskt eftersom sannolikheten för att två förändringar skulle ta ut varandra och att samma checksumma skulle genereras är otroligt osannolik (Giaretta 2011, s. 222).

[48] Även ompaketering och i viss mån refreshment kan påverka autenticiteten hos ett digitalt objekt eftersom dessa migreringar kan påverka proveniensen, kontexten eller dokumentationen kring bevarandet.

[49] Algoritmerna fungerar i det här fallet som komplexa matematiska formler som ger värden till olika delar i de digitala objekten. Algoritmen för att beräkna checksummevärdet hos en textfil exempelvis kan vara skriven så att bokstäverna i alfabetet får olika värden beroende på vilken bokstaven är men också beroende på vilken plats den har i ordet, vilken plats ordet har i meningen eller på sidan och så vidare. En fil med innehållet AAABBB kommer alltså inte att få samma checksumma som en fil som innehåller ABABAB.

Checksummor kan bara visa att ett digitalt objekt inte förändrats så länge det inte transformerats eftersom ett nytt format dels kommer att medföra ett nytt värde på grund av att vissa delar av värdet är knutet till filformatet, men också för att andra förändringar, till exempel rörande strukturen eller utseendet, kan uppkomma.

Förutom att sätta checksummor till enskilda filer kan de även sättas på strukturer som innehåller flera filer vilket kan möjliggöra att man endast behöver kontrollera en checksumma istället för tusentals som ingår i samma struktur, till exempel en mapp eller ett arkivpaket (Adam 2010, ss. 600ff). Dock kan det uppstå problem om den nya checksumman inte överensstämmer med den gamla eftersom man inte kan "räkna baklänges" för att se vilka förändringar som skett i det digitala objektet. Dessutom är en möjlighet att ingenting förändrats i det digitala objektet, utan att den ursprungliga checksumman inte bevarats korrekt och alltså inte är den checksumma som en gång genererades. För att råda bot på detta problem kan flera kopior av checksummorna sparas på olika platser eller i olika system och då de digitala objekten ska kontrolleras jämförs dessa kopior av checksummorna med varandra.[50] Denna metod är utarbetad efter samma mall som LOCKSS,[51] en förkortning för det fantasirika namnet Lots Of Copies Keep Stuff Safe, som inom den digitala biblioteksfären används bland annat för att kontrollera att förändringar inte uppstått i elektroniskt publicerade tidsskrifter och andra digitala biblioteksresurser (Adam 2010, ss. 600ff).

Formatvalidering är en annan metod som kan kopplas till autenticitet hos digitala objekt (Adam 2010, s. 601). Program för formatvalidering kan vara uppbyggda på olika sätt, men principen är att de kontrollerar huruvida ett digitalt objekt är i det filformat som den utger sig för att vara. Detta kan ske mer eller mindre sofistikerat beroende på formatvalideringsalgoritmen, i exempelvis JHOVE, som utarbetats av Harvards universitetsbibliotek och JSTOR, kontrolleras vilka filformat som anges i metadata kring det digitala objektet[52] (Ibid.). Algoritmen i DROID[53] däremot söker genom det digitala objektets bitström och jämför dess egenskaper med kända filformats för att på så vis kontrollera det digitala objektets filformat (Brown 2006, ss. 13ff). Fördelen med formatvalidering kan vara att vid en leverans till ett arkiv kan man kontrollera om de digitala objekten är i det filformat de utges för att vara, vilket kan vara en fördel vid leveranser på tusentals dokument. Dock säger tekniken ingenting om huruvida innehållet i de digitala objekten är det som det uppges vara.

Duranti (2010, s. 81) menar att det ur ett diplomatiskt perspektiv underlättar autenticitetsbevarandet om vissa standardiserade arkivformat används. Vissa filformat

---

[50] Tanken är att om fem kopior av checksummorna är sparade och vid en kontroll är fyra av dessa identiska kan man utröna att något har skett med den checksumma som avviker istället istället för att misstro autenticiteten och integriteten hos det digitala objektet.

[51] http://www.lockss.org/ Lots of Copies Keep Stuff Safe officiell hemsida

[52] http://jhove.sourceforge.net/ JHOVE - JSTOR/Harvard Object Validation Environment

[53] Digital Record Object Identification, är en formatvalideringsalgoritm utarbetad av engelska nationalarkivet i anslutning till deras fria resurs PRONOM, ett register över alla filtyper och versioner av dessa som används globalt. I ett filformat finns vissa regler för hur bitströmmen ska se ut, och dessa katalogiseras för respektive filformat och version i PRONOM (Brown 2006, ss. 7ff) även http://www.nationalarchives.gov.uk/PRONOM/Default.aspx#

har blivit antagna som nationella eller internationella arkivformatsstandarder. RA-FS 2009:2 anger de svenska standardarkivformaten för flera olika typer av digitala objekt, från bilder och dokument till databaser och metadata. Att standardisera arkivfilformaten är ett resultat av att vissa filformat anses lättare att bevara då de är mer "stabila." Bevarandeformatet PDF/A anses exempelvis bättre för lagring av textbaserade dokument än ordbehandlingsformat såsom Word eftersom det inte är lika känsligt för versionsförändringar i programvara samtidigt som det håller den fysiska och logiska strukturen intakt. Det är också ett utslag i att standardisering till vissa utbytesformat underlättar överförandet av arkivmaterial till de statliga arkivinstitutionerna. Enligt Nilsson och Hägerfors (2007, s. 8) kan valet av arkivfilformat påverka den framtida funktionaliteten hos arkivmaterialet. Digitala objekt som endast innehåller text kan både sparas i PDF/A som håller den logiska och fysiska strukturen intakt eftersom representationen lagras likt en bild, eller som textfiler där endast den logiska strukturen bevaras, likt de "enkla kopior" som endast är transkriptioner av originalet (Dollar 2000, s. 28). Om ett digitalt objekt lagras som en textfil underlättas bearbetningen av datan, exempelvis för statistiska ändamål, samtidigt som den fysiska integriteten i någon mening går förlorad, medan PDF/A håller denna intakt samtidigt som den försvårar databearbetning (Nilsson & Hägerfors 2007, s. 8).

Eftersom filformat är konstruerade på olika sätt, antingen genom att metadata bäddas in i filen, såsom för PDF/A, eller att de endast är långa teckenkodade bitströmmar utan inbäddad metadata, som gällande rena textfiler, blir formatmigreringar, det vill säga transformeringar, olika vanskliga. Komplexare format riskerar i högre utsträckning att förändras vid migreringar på sätt som kan få konsekvenser för deras informationsvärde än enklare format. Detta sammantaget med de potentiella skillnaderna i funktionalitet som nämndes ovan gör att vilket arkivbevarandeformat som väljs blir kopplat till det digitala objektets autenticitet eftersom det digitala objektets integritet kan påverkas.

Digitala signaturer diskuteras också som en teknik inom det digitala bevarandet (McNeil 1998, s. 112; Boudrez 2007, ss. 179ff; Gladney 2007, ss. 167ff; Adam 2010, ss. 601–602). Processen sker som följer: Vid en leverans av ett digitalt objekt till ett arkiv skapar avsändaren en checksumma som krypteras, det är denna krypterade checksumma som är den digitala signaturen. Den krypterade checksumman kan endast låsas upp och läsas med hjälp av en så kallad privat nyckel. Den privata nyckeln innehas av den som äger det digitala objektet, till exempel skaparen och inte det arkiv i vilket det digitala objektet förvaras. Detta gör att förändringar inte kan ske i det digitala objektet utan ägarens vetskap, vilket alltså gör att digitala signaturer fungerar ungefär som forna tiders sigill (Duranti 2010, s. 82). När det digitala objektet krypteras skapas även en offentlig nyckel som inte kan dekryptera checksumman, men som används som ett bevis på krypteringen och krypteringsprocessen. Denna nyckel har arkivet och dess användare tillgång till och tanken är att den ger autenticitet till det digitala objektet genom att den kan garantera den digitala signaturens riktighet även om den inte kan öppna krypteringen. I likhet med checksummor fungerar inte digitala signaturer i de fall det digitala objektet måste migreras genom transformation av formatet eftersom den ursprungliga checksumman då oundvikligen inte kommer att gälla längre (Boudrez 2007, s. 187).

Proveniens är ett centralt begrepp i arkivvetenskapen. Med hjälp av proveniensen skapas det kontextuella sammanhang som Hirtle (2000, s. 10) menar är förutsättningen för att en institution ska vara ett arkiv och inte ett lager. Enligt Adam (2010, s. 8) är att dokumentera ett digitalt objekts proveniens att "capture the history of authentication," vilket gör att ett digitalt objekt i ett system för långtidsbevarande bör sammankopplas med dokumentation av bevarandeåtgärder. Detta kan ta sig uttryck i metadataschema, loggar och annan dokumentation som innehåller information om de tekniska autenticitetsmarkörerna som beskrivits ovan, tillsammans med tidsstämplar som anger en kronologi och behörighetsinformation som anger vem som utfört åtgärden (Giaretta 2011, ss. 22, 184–187).

Ur ett mer traditionellt arkivvetenskapligt perspektiv har proveniens kopplad till autenticitet beskrivits i teoriavsnittet om diplomatik och kommer därför inte att beskrivas mer utförligt här.

8.1.3 Signifikanta egenskaper och kopior

Levy (2000, s. 37) menar att den digitala världen endast består av oändliga kopior eftersom "samma" dokument kan se olika ut vid olika framställningar då det digitala dokumentets fysiska struktur är så instabilt. Ett tankeexperiment kan förtydliga: ett dokument sparat på en hårddisk återges på datorskärmen på ett visst sätt när det skapas. Precis efter att dokumentet färdigställts går datorskärmen sönder och en ny skärm med annan upplösning och andra färginställningar införskaffas. Även om ingenting i det digitala objektet förändrats så har skillnader i hur originaldokumentet uppfattas av användaren uppstått.

Det stora problemet kring vad som är ett original infinner sig då digitala objekt migreras från ett filformat till ett annat eftersom detta förändrar bitströmmen. Bitströmmen är det som många bevarandestrategier faller tillbaka på som originalet eftersom det ses som ett surrogat för dokumentet (Gladney 2007, s. 153). Gladney menar vidare att detta inte håller konceptuellt eftersom samma bitström kan ge olika representationer beroende på vilken mjuk- och/eller hårdvara som används för att återge den. Dessutom innebär ett synsätt där bitströmmen är originalet att ett digitalt objekt som migrerats till ett nytt format för att behålla sin läsbarhet inte längre är i original.

Om inte bitströmmen är originalet, finns då andra alternativ? Hänström (2007, s. 85) menar att det finns en internationell överenskommelse att det är representationen som ska bevaras, vilket återknyter till performancemodellens syn på digitalt bevarande (Heslop et al. 2002, ss. 8–10). Detta kan dock i praktiken vara problematiskt eftersom det inte finns något sätt att jämföra huruvida två bitströmmar i olika filformat men av samma digitala objekt har bevarat samma representation utan att jämföra bitströmmarnas representationer visuellt med varandra (Gladney 2007, s. 153). Dessutom kan man fråga sig hur lika två representationer måste vara för att anses återge samma information till betraktaren, och om betraktaren själv har någon påverkan på vilka likheter som måste vara uppfyllda. Innan dessa frågor besvaras bör man tänka över vilken funktion bevarandet av digitala objekt fyller, nämligen att de ska bevara *information* om någonting, bevarandet av de digitala objekten i ursprunglig form är inte ett självändamål i sig.

För att komma tillrätta med denna typ av frågeställningar har man skapat en teoribildning kring hur representationer och dokument kan beskrivas och kvantifieras i informationsbärande enheter som kallas *significant properties*, på svenska signifikanta egenskaper. Dessa har även benämnts *salient features* och *essential characteristics* även om konceptet har varit detsamma enligt Yeo (2010, s. 87). Begreppet signifikanta egenskaper har under årens lopp fått något olika definitioner, vilket delvis kan förklaras av att samma uttryck används av forskningsområden med liknande, men inte helt överensstämmande, mål och utgångspunkter (Faniel & Yakel 2011, ss. 156–159). De forskare och projekt som arbetat med signifikanta egenskaper inom digitalt bevarande utgår från en data-centrisk syn där begreppet får symbolisera de egenskaper hos det digitala objektet som behövs för att det ska vara läsbart och för att informationsinnehållet ska vara intakt. Detta ställs mot de som arbetar med dataåteranvändning (data reuse) som ofta tar ett etnografiskt perspektiv som ställer användarens bruksmöjligheter i centrum och där man menar att signifikanta egenskaper är de faktorer som förvaltar ett objekts mening över tid och gentemot en målgrupp (Ibid.). Då detta jämförs med den infologiska ekvationen kan man ana att de forskare som arbetar med digitalt bevarande lägger störst tyngd vid datavariabeln medan förförståelsevariabeln, som ju är beroende av användargruppen, är den som främst behandlas av de som forskar på dataåteranvändning. Oavsett forskningsfokus är den minsta gemensamma nämnaren att signifikanta egenskaper är aspekter hos det digitala objektet som krävs för att dess datainnehåll ska kunna användas som information över tid.[54]

Nilsson (2008, ss. 26–27) menar att digitala objekts fysiska sammansättning, så som de presenteras visuellt på en datorskärm eller i en utskrift, kan översättas i metadataelement. Genom att göra detta på ett sätt som möjliggör för framtida återläsning av objekten i en form som påminner om den ursprungliga kan ett digitalt objekts signifikanta egenskaper bevaras. Giaretta (2011, ss. 216–217) betonar att signifikanta egenskaper är beroende av målgruppens behov: för en målgrupp med en viss förförståelse och vissa behov kan det vara av vikt att behålla textkulören i ett dokument vilket gör färgen till en signifikant egenskap, samtidigt som det kan vara onödigt för en annan målgrupp, varpå den inte är en signifikant egenskap för denna. Eftersom inte alla aspekter hos ett digitalt objekt anses vara signifikanta egenskaper och då dessa är beroende av målgrupp, faller det på arkivarien att bestämma vilka egenskaper hos ett digitalt objekt eller handling som bör bevaras (Nilsson och Hägerfors 2007, s. 9). Detta då arkivarien är specialist på bevarande och tillgängliggörande av information, och att välja bort egenskaper hos ett digitalt objekt som inte ska bevaras kan likställas med gallring (Ibid.). För att citera Eastwood (1994, s. 125)

[54] Det engelska begreppet significant properties är således inte helt enkelt att översätta till svenska eftersom vissa aspekter av begreppet kommer att framhävas eller åsidosättas beroende på ordvalet. Som diskuteras används begreppet av flera forskningsdiscipliner och har därför skilda betydelser för dessa, vilket måste tas i beaktande vid en översättning. Valet att översätta begreppet till signifikanta egenskaper gjordes för att poängtera att begreppet i denna uppsats snarare används i den datacentriska synen där informationsinnehållet ska vara intakt än i den etnografiska synen som poängterar meningen hos objektet. Gällande det senare användningsområdet hade signifikanta eller essentiella karakteristika varit att föredra.

> The archival discipline consists in building knowledge about archival documents and acting upon them in methodical ways to protect the properties that they have. Thus, the large theoretical question is what are those properties that need to be protected and why?

Vilka egenskaper hos ett digitalt objekt eller ett dokument kan beskrivas med signifikanta egenskaper? Ett dokument kan sägas ha en logisk och en fysisk struktur, där den logiska strukturen är hur de intellektuella elementen relaterar till varandra, exempelvis i vilken ordning rubriker och kapitelindelning förekommer, och den fysiska beskriver layout, fonter, färger och så vidare (Nilsson och Hägerfors 2007, s. 4). Som tidigare konstaterats kan båda dessa strukturer, beroende på målgruppens behov, utgöra signifikanta egenskaper. Samma idéer återkommer inom diplomatiken, där man studerar ett dokuments *intrinsic* (motsvarande logisk) och *extrinsic* (att jämföra med fysisk struktur) element för att avgöra autenticiteten hos detta (Duranti 1998, s. 151).

Signifikanta egenskaper kopplade till logisk struktur har studerats mer utförligt än de kopplade till fysisk struktur, särskilt inom biblioteks- och informationsvetenskap (Yeo 2010, ss. 89–90). Den logiska strukturen har tidigare setts som viktigare att bevara än den fysiska, vilken dock av Yeos forskningssammanställning att döma, har börjat få en renässans. Att den fysiska strukturen varit så styvmoderligt behandlad kan bero på att det behövs mycket metadata för att beskriva hur digitala objekts fysiska struktur ska representeras visuellt på en datorskärm eller i en utskrift (Nilsson 2008, s. 41). Detta gäller särskilt då objektet ska bevaras under en längre tid och den fysiska strukturen ska kunna återges efter ett antal migreringar, och då den återgivande mjukvaran inte är den samma som en gång användes för att återläsa och representera bitströmmen.

Då metadata kan användas för att beskriva och bevara signifikanta egenskaper knutna till de logiska och fysiska strukturerna i det digitala objektet kan hela representationen beskrivas för att kunna återges så oförändrat som möjligt (Nilsson 2008, ss. 26–27). Alternativt, i de fall alla egenskaper hos det digitala objektet eller dokumentet inte måste bevaras, kan man bevara signifikanta egenskaper som gör att man skapar vad som kan anses vara en kopia av originalet. Att egenskaper hos informationsbärande enheter som dokument och handlingar kan förändras eller förloras är egentligen ingenting nytt för det digitala samhället, problemet har funnits så länge man kunnat göra kopior. Signifikanta egenskaper, och de avvägningar som är förknippade med att välja ut dessa, kan förekomma vid alla konverteringar från ett medium till ett annat, till exempel från papper till microfilm (Yeo 2010, s. 87). Detta gäller i de fall konverteringen innebär att det objekt som konverteras kan få förändrade egenskaper, exempelvis om ett flerfärgsoriginal konverteras till ett monokromt medium eller förminskas så att man inte kan utläsa detaljer.

Eftersom potentiella signifikanta egenskaper riskerar att falla bort vid en formatmigrering kan man inte anse att det migrerade digitala objektet ens är en exakt kopia (Adam 2010, s. 597) eller "kopia i samma form som originalet" för att återknyta till diplomatiken (Dollar 2000, s. 28). Om objektet är läsbart efter migreringen kan det istället ha blivit en "imiterande" eller "enkel" kopia vilket kan ha påverkat dess informationsinnehåll för vissa målgrupper enligt den infologiska ekvationen.

Nilssons kvalitativa undersökningar vid svenska arkivinstitutioner har visat en tendens till att arkivarier prioriterar metadata om proveniens och kontext framför metadata om handlingarnas fysiska utseende (2008, ss. 34–35). Detta gör att de autenticitetsaspekter som kan utläsas via proveniens- och kontextbevarande metadata kommer att utgöra merparten av den information om ett digitalt objekts autenticitet som en framtida användare kan ta ställning till. Viktigt att betänka i sammanhanget är att det som inte dokumenteras inte kan bevisas; om det ursprungliga utseendet hos ett digitalt objekt så som det såg ut på datorskärmen aldrig dokumenterats och inte heller de förändringar som skett på grund av formatmigreringar och dylikt kan framtida användare inte ta ställning till om det digitala objektets integritet är intakt.

8.1.4 Teoretisk belysning av litteraturöversikten

Innan intervjuresultaten redovisas kan det vara lämpligt att återknyta till uppsatsens definition av autenticitet. *Ett digitalt objekt är autentiskt om det digitala objektets integritet och proveniens/identitet säkrats och kan bevisas.* Bibehållen integritet innebär att objektet inte blivit korrumperat avsiktligt eller oavsiktligt, och att eventuella förändringar av integriteten under exempelvis migreringar noga dokumenterats. Vidare är det avgörande att läsbarheten har bibehållits. Proveniens kommer ur klassisk arkivteori och diplomatik och har ovan beskrivits som att säkerställa det digitala objektets identitet, ursprung och skapandekontext. Från diplomatisk teori kommer även tanken om att kunna visa ansvarskedjan. Om integritetsdokumentation visar vad som hänt med dokumentet visar ansvarskedjan vilken institution som haft hand om dokumenten, vilket i sin tur kan knytas till tanken om tredjepartsintygandet som garant för autenticitet.

Sålunda innehåller integritetsbegreppet i uppsatsens definition autenticitetsmarkörerna checksummor, signifikanta egenskaper, formatvalidering, digitala signaturer och dokumentation om händelser relaterade till dessa. Proveniensbegreppet i sin tur markeras av identitet, ursprung, kontext (processer och arkivbeskrivningar till exempel), samt dokumentation om händelser relaterade till dessa. Ansvarskedjan placeras i diplomatisk teori under integritetsbegreppet, det har dock även en relation till proveniensens ursprungstanke. Tillit och tredjepartsintygande är även viktiga begrepp kopplade till autenticitet. Dessa rör sig dock ovanför objekt- och dokumentationsnivån och är mer sociologiska och psykologiska till sin karaktär.

Figur 15. Tabell över resultatet av litteraturgranskningen kopplat till det teoretiska perspektivet.

Autenticitetsmarkör	**Syfte**	**Autenticitets-aspekt**	**Härleds från teoretiskt perspektiv**
Checksummor	Dokumenterar bitströmmens oföränderlighet	Integritet	OAIS (förändringsskydds-PDI)
Formatvalidering	Kontrollerar och dokumenterar filformatet	Integritet	OAIS (representations-information)
Digitala signaturer	Dokumenterar bitströmmens oföränderlighet, krypterat	Integritet	OAIS (förändingsskydds-PDI), diplomatik
Identitet	Namnger objektet med unik identifikator	Proveniens	OAIS (referens-PDI), diplomatik
Ursprung	Härleder bland annat objektet till arkivbildaren	Proveniens	OAIS (proveniens-PDI), diplomatik
Kontext	Dokumenterar objektets skapandekontext	Proveniens	OAIS (sammanhangs-PDI, diplomatik
Signifikanta egenskaper	Dokumenterar informationsbärande egenskaper hos objektet	Integritet	Diplomatik
Ansvarskedja	Dokumenterar bevarandet av objektet	Integritet, proveniens	Diplomatik, OAIS (proveniens-PDI)

8.2 Information eller handling?

Anser informanterna att de bevarar handlingar eller information, och hur ser de på originalitet hos elektroniska dokument.

Alla informanterna är eniga att det är informationsinnehållet snarare än handlingarnas ursprungliga bitström som ska bevaras. När det gäller originalbegreppet i digitala sammanhang har en del av informanterna övergett begreppet "original" till förmån för "förlaga". Glidningen i betydelse innebär att det är bevarandet av informationens originalitet som det digitala bevarandet lägger fokus på, snarare än handlingen i sig.

En anpassning bort från fysiska definitioner av originalitet sålunda. Informanterna menar att detta är en nödvändig följd av de förändrade förutsättningar som den digitala världen för med sig.[55] Resonemanget blir väldigt tydligt när efter digitalisering av analogt material det analoga dör eller gallras. Kopian tar då över rollen som original, påpekar informant 6 från Riksarkivet när denne filosoferar om ämnet:

> Ja, och vad är det som ska vara samma. Det är aldrig samma. Vad är det viktiga liksom att bevara? Där har vi ju, om man gör en migrering från analogt till digitalt till exempel, då blir det ju en kopia på det analoga som vi gör, men den analoga dör kanske, så att den inte finns längre. Och då blir det kanske den digitala som blir originalet. Men hur vet du att den är originalet?

Om sättet att tänka kring original har förändrats hos de som arbetar med e-arkiv, tar några informanter upp frågan att en viss tröghet i anpassning till detta kvarstår hos arkivbildar- och användargrupper. Exempelvis berättar informant 1 vid Stockholms stadsarkiv om dennes diskussioner med jurister som menat att pappersoriginalen av dokument som digitaliserats bör bevaras då dessa är underskrivna, eftersom de menar att dessa signaturer skapar en särskild äkthet. Informanten själv menar att de skannade exemplarens autenticitet inte behöver vara sämre än pappersförlagorna då det finns "processer som säkerställer att de inte kommer att förfalskas, på samma sätt som man säkerställer att ett pappersdokument inte ligger framme för obehöriga att komma åt att förfalska.”

8.3 Metoder och autenticitet

8.3.1 Autenticitetsmarkörer som används av institutionerna

Vilka metoder används av institutionerna för att bevara autenticitet och vilka har man valt bort?

Samtliga arkivinstitutioner har valt att använda sig av migreringsstrategin, men ingen av dem har ännu ställts inför en formatmigrering av arkivbevarandeformat. Resultaten över vilka metoder de använder och hur de implementerar dessa presenteras i en tabell nedan.

[55] Dock finns det försök att via den så kallade performancemodellen kunna behålla iallafall en visuell originalitet (Se avsnitt 5.1).

Figur 16. Tabell över autenticitetsmarkörernas användning på institutionerna.

Autenticitets-markör/metod	**Riksarkivet**	**SYLL**	**Stockholms stadsarkiv**
Checksummor	På de enskilda objekten och arkivpaketen	På de enskilda objekten	På de enskilda objekten
Digitala signaturer	Bevaras ej	Bevaras ej	Bevaras ej
Formatvalidering	Sker när objektet tas in i systemet	Ej utrett	Sker när objektet tas in i systemet
Proveniens: Identitet, ursprung och kontext	Identitetsmarkör, arkivbildare, arkivbeskrivning, arkivförteckning, annan kontextskapande dokumentation.	Identitetsmarkör, arkivbildare, arkivbeskrivning, arkivförteckning, annan kontextskapande dokumentation.	Identitetsmarkör, arkivbildare, arkivbeskrivning, arkivförteckning, annan kontextskapande dokumentation.
Ansvarskedja	Dokumenteras i metadata.	Dokumenteras i metadata.	Dokumenteras i metadata.
Loggar över ansvarskedjan skapade hos arkivbildarna	Bevaras ej	Bevaras ej	Bevaras ej
Signifikanta egenskaper	Kan bli aktuellt vid migrering, i nuläget dokumenteras original representation av levererande myndighet. Handböcker och systemmanualer över de skapande systemen bevaras.	Kan bli aktuellt vid migrering, i nuläget dokumenteras original representation av levererande myndighet.	Kan bli aktuellt vid migrering, i nuläget dokumenteras original representation av levererande myndighet.

Checksummor

Genom att kontrollera att checksummorna överensstämmer varje gång ett digitalt objekt flyttats i arkivsystemet eller mellan olika databärare kan man kontrollera att ingenting har förändrats i objektet i dessa överföringar enligt informant 3. RA har planer att spara paketens checksummor på tre separata fysiska platser berättar informant 5. Detta är en del av LOCKSS-metodologin, att spara digitalt material på flera olika ställen.

Digitala signaturer

Ingen av institutionerna bevarar digitala signaturer. På RA menar informant 5 att digitala signaturer inte är lämpliga för långtidsbevarande då de ej går att migrera och är beroende av så kallade rekursioner som endast är valida några år, vilket gör signaturerna väldigt tidsberoende. Syftet med dessa verkar vara mest för förvaltningarna själva. De är inte relevanta för arkivet, säger informant 6. Ett sätt att uttrycka svårigheten med digitala signaturer i ett långtidsperspektiv är att krypteringen kan vara väldigt vansklig att bevara, migrera och autenticera över tid, men i och med att filerna lagras dekrypterade kan man ju ändå kontrollera att inget hänt med dem genom att använda checksummor, menar informant 4.

Formatvalidering
Ingen av de undersökta institutionerna använder organiserad formatvalidering. SYLL planerar att ta emot alla format och sedan ändra till det format de vill ha, menar informanten. Processen har dock inte startat upp än inom SYLL, vilket gör att frågan ej kan besvaras slutgiltigt där. Det är vidare inte aktuellt för de andra institutionerna då de bara tar emot vissa filformat som regleras i RA-FS 2009:2, innan leverans ska filerna ha konverteras till dessa arkivformat. Formatvalidering i arkivteoretiska sammanhang har mest relevans i länder där de offentliga arkiven inte kravställer de format som levereras, utan vad som helst kan dyka upp vid leveranser. I USA är det exempelvis nödvändigt att formatvalidera digitala arkivleveranser av denna anledning, upplyser informant 6 om.[56]

Samtliga studerade arkivinstitutioner bevarar eller ämnar att bevara de digitala objekten i vissa förutbestämda filformat. De bevarandeformat som valts av RA och SSA är sådana som stipuleras i RA-FS 2009:2, exempelvis PDF/A, TIFF, JPEG, XML och HTML (RA-FS 2009:2 kap. 3). I RA-FS 2009:2 finns inga filformat angivna för ljud och video. SSA och RA har upplevt att trots att man använder PDF/A som är en ISO-standard har det hänt att digitala objekt i PDF/A-format har "studsat" när de lagts in i arkivet. Detta har visat sig bero på att ISO-standarden för PDF/A är tämligen vid och öppen, vilket innebär att PDF/A-formatet hos två digitala objekt kan skilja sig från varandra om de konverterats eller skapats med olika program. Detta har lett till att man slagit fast vilken nivå på standarden som ska följas, så att PDF/A-filerna ska ha samma minsta gemensamma nämnare som arkivplattformen godkänner. För att underlätta för stadens förvaltningar har man köpt in en gemensam PDF/A-konverterare som konverterar till denna specificerade nivå. SYLL ämnar bevara de digitala objekten i vissa format, men då de kommer att förvalta digitala objekt som är skapade i många olika system och filformat planerar man att införskaffa konverterare så att dessa kan föras över i arkivformaten. I dagsläget har man endast konverterare mellan ordbehandlingsdokument och PDF/A, men efterhand som leveranser tas emot och man ser vilka filformat man har behov av att konvertera kommer man att införskaffa fler.

Beroende på vilket format det digitala objektet sparas i får man olika funktionalitet, på grund av detta har RA valt att spara hemsidor i HTML så att navigeringen mellan olika delar av sidan kan bevaras vilket även bevarar och förtydligar kontext och samband. Detta till skillnad mot om man sparat i PDF vilket gjort att dessa funktioner försvunnit, samtidigt som det möjligtvis varit lättare att bevara. Samma frågeställningar kring skillnader i versioner och formatvarianter har dykt upp gällande digitala objekt i TIFF-format eftersom detta också är ett brett och vildvuxet filformat, menar informant 6.

Även om informanterna uppger att institutionerna inte använder formatvaliderare som verktyg i det digitala bevarandet, sker trots allt en formatvalidering av de digitala

[56] Vilket stöds av information från Library of Congress (2008).

objekten då de läggs in i e-arkivplattformen eftersom de då kan "studsa" i de fall filformatet inte lever upp till den minsta gemensamma nämnaren.

Proveniens/identitet
Institutionerna använder metadatascheman och metadatastandarder för att beskriva proveniensrelaterade autenticitetsmarkörer. Dessa är arkivbeskrivningar, arkivförteckningar och de processbeskrivningar som sätter arkivmaterialet i en kontext. SSA beskriver proveniensmetadata i två scheman, ett kring organisationen och ett kring arkivobjekten berättar informant 1.

Att dokumentera det egna arbetet
Att säkra ansvarskedjan och i samband med detta på ett tydligt sätt dokumentera rutiner och arbetssätt på arkivet bedöms vara viktigt av samtliga informanter. Det centrala är att och hur man dokumenterar bevarandet genom att dokumentera arbetsrutiner, hur organisationen sett ut, vilka handböcker man haft och vad som stått i dem, vilka checklistor man haft för leveranser och hur den tekniska miljön sett ut. Att kunna visa och beskriva detta och kombinera exempelvis med checksummor ger en bra helhetsbild, menar informant 1.

> Och det blir kanske i två led, både hur informationen har hanterats ute i verksamheterna/.../Och att vi har skött det på ett schysst sätt under den tiden vi har haft det hos oss. Där igen, alltså tekniken skiljer ju naturligtvis, men det är litegrann samma sak som i analoga handlingar, att det är samma, på något vis, frågeställningar om man väljer att lita på, det är kanske svårt att veta med 100 procents säkerhet att en 100 år gammal handling är det korrekta.
>
> *Informant 1*

Förutom att bevarandehistorik visar ansvarskedjan kan den även visa att arkivet arbetat enligt så kallad *best practice* påpekar informant 6. Detta innebär att man visar för framtida användare att arkivet arbetat enligt de vedertagna metoder som för tiden funnits. Att kunna visa att man arbetat enligt *best practice* är en bra metod för att öka tilliten till materialet ur ett långtidsperspektiv menar även informant 1. Informanten poängterar vidare att ett viktigt sätt att göra detta är att arbeta proaktivt med det digitala materialet och inom bevarandeplaneringsfunktionen i OAIS-modellen blicka framåt för att kunna hålla sig uppdaterade kring de senaste förutsättningarna för digitalt bevarande gällande tekniska lösningar, filformat och så vidare.

Loggar från myndigheter och förvaltningar
I vissa ärendehanteringssystem skapas automatisk metadata varje gång en ändring sker i ett digitalt objekt, vilket sparas i loggar, de beskriver alltså allt som skett med dokumentet när det fortfarande fanns hos arkivbildarna och inte än nått arkivet. Inga av de undersökta arkivinstitutionerna är dock intresserade av att bevara dessa. RA menar att loggarna skapar förvirring då de innehåller så mycket information att de blir ohanterliga, de är ofta flerfaldigt större än de allmänna handlingar som ska bevaras, säger informant 6. Dessutom menar våra informanter att det är relativt ointressant information. "[...]Men om man skickar med hela loggen. Man ser inte träden för skogen eller skogen för träden liksom, det går inte. Det är ingen idé. Loggar kan stjälpa lika mkt som hjälpa", påpekar informant 6. Myndigheter rensar även ofta ut loggarna själva för det blir så tungrott, och ett sätt att se på loggar är som arbetsmaterial som hade gallrats i att analogt system också, menar informant 5. På

SSA följer Ioggar kring vem som skrivit i ärendehanterings-, socialtjänst- och andra verksamhetssystem inom kommunen i normalfallet inte med. Informanten ser dock en öppning för, beroende på hur loggningsfunktionaliteten i systemen ser ut, att det finns en möjlighet att loggarna kunde vara relevanta att bevara. Ofta finns det dock enligt informanten en del processer på förvaltningarna som gör bevarande av loggar överflödiga:

> /.../det här med att när man tar fram dokument i ett verksamhetssystem då, ute i verksamheten, så ansvarar myndigheten för att göra det på ett korrekt sätt. Och på ett säkert sätt, så att man har behörighetsstyrning inne i systemen och man har liksom olika sätt som gör att det inte ska gå att hacka sig in, och man har back-up. Då ska man liksom, utifrån att man har korrekta processer kunna lita på att den dokumentation som finns i systemen är korrekt.
>
> *Informant 1*

Inom SYLL har man inte heller intresse av att bevara loggar, uppger informant 2.

Att bevara integriteten vid formatmigreringar

Ingen av de undersökta institutionerna har hittills stått inför att genomföra formatmigreringar mellan olika standardarkivformat. På grund av detta är svaren kring detta främst resonerande och spekulationer eftersom institutionerna har valt att inte formalisera rutiner kring formatmigreringar förrän de ställs inför dem. Detta då det bedömts omöjligt att göra avväganden kring problemet utan att känna till under vilka förhållanden formatmigreringen kan ske enligt informant 3 och 6.

Riksarkivet har genomfört migreringar av gamla textfiler där man bytt teckenkod från EBCDIC till ASCII, som man i stort sett kan se som identiska filer då den logiska och fysiska strukturen av filen bevarats, liksom tillhörande metadata[57], det enda som förändrats är i vilken teckenkod siffrorna och bokstäverna är skrivna i. Denna typ av migrering är enklare att genomföra utan att potentiellt sett förlora eller förändra information än formatmigreringar av mer komplexa och sammansatta filformat som exempelvis PDF, där metadata finns "inbakad" i det digitala objektet via filformatet. Vilka förändringar som sker beror dock på vilket konverteringsverktyg som används och hur detta programmeras. Informant 6 exemplifierar med att man vid en formatmigrering från PDF till PDF/A måste ha alla de typsnitt som PDF/A-filen ska innehålla sparade i datorn som genomför migreringen eftersom denna annars inte kan inlemma dessa i den migrerade filen, varpå migreringen inte kan genomföras med det tänkta resultatet. I de fall Riksarkivet behövt migrera digitala objekt från PDF till PDF/A har metadataförändringarna dokumenterats i tjänsteanteckningar snarare än som metadata relaterad till de digitala objekten.

Inför en formatmigrering, resonerar informant 6, bör en noga dokumenterad riskanalys genomföras både gällande tekniska faktorer och signifikanta egenskaper. Sedan måste själva migreringen och de eventuella förändringar som uppstår

[57] Det som förändrats är hur många ettor och nollor och i vilken sammansättning dessa har, men ett A tolkas och representeras fortfarande som ett A, bara datorn får information om vilken teckenkodning den ska använda för att tolka bitströmmen.

dokumenteras i bevarandemetadatan. Det kan vara aktuellt att göra stickprov för att kontrollera att migreringen genomförts enligt planerna och att representationerna bevarats resonerar informant 3. Digitala objekt som förvaras på två olika fysiska platser bör migreras på samma sätt menar informant 4. För att ha möjlighet att göra om migreringar i framtiden om det visar sig att materialet korrumperats på något sätt, antingen gällande autenticitet eller andra faktorer, bevarar Riksarkivet även de omigrerade kopiorna. Huruvida det blir aktuellt att bevara alla mellanled i migreringsprocessen är oklart i nuläget eftersom det är svårt att sia om hur stora datamängder det kommer att röra sig och även vilka förändringar som skett hos de digitala objekten i och med migreringarna.

Även SSA står inför formatmigreringar från PDF till PDF/A eftersom digitala objekt i vissa e-leveranser som inkommit till arkivet innan e-arkivet togs i bruk var i detta filformat. Informant 1 uppger att man inom arkivinsitutionen resonerat kring att det omigrerade digitala objektet, vilken man i någon mån kan se som originalet eller förlagan, inte behöver bevaras efter en migrering eftersom den migrerade kopian ska vara tillräckligt bra för att ersätta förlagan. Detta sätt att resonera använder sig arkivinstitutionen av även då pappersförlagor digitaliseras då originalen gallras efter scanning. Informant 1 resonerar kring att vid en formatmigrering är det viktigt att utreda vilka signifikanta egenskaper som kan finnas i de digitala objekten och hur man ska gå tillväga för att bevara dessa. Informanten exemplifierar med att i planritningar är det oumbärligt att färgen bevaras eftersom man annars inte kan skilja olika ledningstyper från varandra. Mindre meningsbärande element skulle man dock kunna fatta gallringsbeslut kring i de fall de svårligen kan inkluderas som signifikanta egenskaper filosoferar informanten vidare.

Hur formatmigreringar ska hanteras är en av de frågor som man inom SYLL ska börja arbeta med och ta ställning till. Det som hittills sagts är att varje digitalt objekt ska bevaras i det format det inkom i till bevarandesystemet, samt det arkivformat till vilket det migreras. Om alla mellanled ska bevaras har man ännu inte tagit ställning till.

8.3.2 De viktigaste autenticitetsmarkörerna

Vilka markörer ser informanterna som avgörande för ett digitalt objekts autenticitet i ett långtidsperspektiv?

Informanterna tillfrågades om vad de själva tyckte var den viktigaste markören för att bevara autenticiteten hos digitala objekt. Checksummor ses av samtliga informanter som avgörande för digitala objekts autenticitet, antingen allena eller i kombination med andra markörer. Checksummeomröstningar kan vara ett viktigt komplement till dessa enligt informant 5, samt att ha rutiner för checksummekontroller varje gång de digitala objekten flyttas mellan databärare eller system menar informant 3.

Informant 1 menar att förutom att kunna visa den tekniska autenticiteten är proveniensmetadata viktig för att skapa en autenticerande kontext för det digitala objektet.

Informanterna 4 och 6 menar att det är viktigt att man kan redovisa hur det digitala bevarandet gått till. Informant 4 poängterar särskilt vikten av att man i

bevarandemetadatascheman såsom PREMIS dokumenterar arkivets handhavande av de digitala objekten, vilka åtgärder som görs, varför de görs och vem som är ansvarig för dem. Informant 6 sammanfattar sin syn på frågan:

> Checksummor och så interna rutiner. Man måste ha koll på hela ledet, från vaggan till graven liksom, det är ju det/.../både i PREMIS [det bevarandemetadataschema som RA använder. Uppsatsförfattarnas kommentar] och i beskrivningar och i all dokumentering kring hela det digitala bevarandet.

8.4 Vilka kommer användarna att vara?

Vilka framtida brukargrupper ser informanterna till det elektroniska materialet och vilka användningsområden och förutsättningar tror de att dessa kommer att ha?

Enligt OAIS-modellen, som samtliga arkivinstitutioner menar att deras e-arkiv är uppbyggda efter, sker digitalt bevarande med en särskild målgrupp, den så kallade designated community, i åtanke. Vilka behov och förkunskaper denna målgrupp har påverkar vilken metadata som väljs ut för bevarandet, eftersom dessa ska skräddarsys efter målgruppen. Då arkivinstitutionernas arkivmaterial innehåller allmänna handlingar är det intressant att undersöka hur institutionerna förhåller sig till framtida målgrupper och deras behov.

Samtliga informanter är eniga om att de allmänna handlingar de bevarar ska kunna göras tillgängliga för allmänheten och att denna därför är en given målgrupp. Förutom denna ser informanterna även andra mer specialiserade brukargrupper. Inom SYLL exempelvis är det rimligt att tänka sig att dokument som skapats inom forskning främst kommer att användas i framtiden av forskare inom den disciplin de skapats, eller andra fält som ligger nära denna. Då kan det vara en styrka för dessa dokument och arkiv att de beskrivs i metadata av den forskargrupp som tog fram dem eftersom de kan materialet bättre än någon annan och även vet vilken ytterligare information som kan behövas för att tolka materialet, resonerar informant 2.

En särskild brukargrupp för SSA identifieras som stadens befolkning som man har ambitionen att fungera som ett informationsnav för. Istället för att medborgarna ska gå via stadsförvaltningarna för att ta del av allmänna handlingar ska dessa levereras tidigt i "sin levnad" till stadsarkivet. Själva metadatan som levereras kring handlingarna behöver inte påverkas av detta, men det kan vara aktuellt att förändra sökfunktionaliteten och representationen av systemen om de ska användas av allmänheten snarare än förvaltningarna som skapade dem, menar informant 1. Informant 4 från Riksarkivet menar att en hel del arkivmaterial används, forskas på och studeras på sätt som de inte ursprungligen skapades för, exempelvis används husförhörslängder som underlag för demografiska undersökningar. Det är därför svårt att kunna föreställa sig idag vilka de framtida användningsområdena kommer att vara för arkiv som levereras nu. På grund av detta menar informanter från Riksarkivet att det bästa är att ta emot så mycket metadata som möjligt i samband med leveranserna så får framtida användare själva sålla bland informationen. Istället för att lägga till rätta leveranserna med en särskild målgrupp i åtanke försöker man bevara och förteckna arkiven så ursprungligt som möjligt och tar emot metadata både av teknisk och kontextuell karaktär. Informant 4 menar vidare att myndigheter lägger till

metadata i sådan utsträckning att den befolkningsgrupp som myndigheten har åtaganden gentemot kan förstå och bruka materialet, exempelvis sjukvården gentemot sjukvårdspersonal, patienter och anhöriga. På så vis blir denna medborgargrupp målgruppen. Samtidigt berättar informanten att denne på arkivkonferenser hört företrädare från forskningen propagera för att mer forskningsspecifik metadata bör läggas till de digitala objekten redan ute på förvaltningarna. Detta ses dock som alltför tidskrävande av företrädare från förvaltningarna som menar att deras uppdrag består i att handlägga ärenden och inte metadata som kanske blir aktuell i framtida forskning.

8.5 Överlämningen till e–arkivet

Hur ser överlämningsprocessen, med särskilt fokus på autenticitet, av digitala objekt till de offentliga arkiven ut?

Vid överlämning till SYLL och SSA måste levererande myndighet, förvaltning eller arkivbildare tillsammans med arkivet arbeta ut en plan för vilka filformat de digitala objekten ska vara i, vilken metadata som ska sättas till, vilken dokumentation kring systemet som de digitala objekten är skapade i som ska ingå och så vidare. Vid leverans till SYLL kallas dessa för inleveransutredningar och till e-arkivet hos SSA benämns de anslutningsprojekt.

Formen för hur inleveransutredningarna ska se ut inom SYLL håller för tillfället på att fastställas genom ett pilotprojekt där digitalt material från ett antal olika lärosäten ska provlevereras. På grund av detta kan denna process inte beskrivas mer detaljerat. Vissa moment som kommer att ingå har redan identifierats; inom SYLL kommer de levererade digitala objekten att konverteras till vissa förutbestämda arkivformat och man kommer att använda ett standardmetadataschema. Detta kommer dock att ha valfria element som vid behov kan läggas till beroende på leveransen och arkivbildarens behov. Förutom metadata om systemen som leveransen genererats ur, kommer såväl automatisk metadata som skapats om de digitala objekten i systemet och deskriptiv metadata för att beskriva leveransen att ingå. Denna beskrivning skapas av den som levererar under vägledning av arkivinstitutionen som mottar leveransen in i SYLL, likaså är valet av valfri metadata och vilken systemdokumentation som bör medfölja något som arbetas fram i inleveransutredningarna. SYLL kommer att lagra de digitala objekten i de format de levereras såväl som konverterade till ett arkivformat. Eftersom det finns en potential att SYLL kommer att få leveranser i många olika, mer eller mindre vanliga, filformat kan det dröja innan en leverans konverteras till ett arkivformat då det kan vara att en särskild konverterare måste tas fram om den innehåller en filtyp som tidigare inte konverterats.

Instruktioner och manualer för hur anslutningsprojekten till e-arkivet vid SSA ska bedrivas finns att ta del av via deras hemsida.[58] Dessa projekt är tämligen tids- och arbetskrävande då de inbegriper ett flertal möten mellan arkivet och levererande myndighet. Vid dessa möten instrueras leverande myndighet eller förvaltning i hur de ska fylla i metadataschemana, hur myndigheten bör beskrivas i de deskriptiva elementen, vilket djup som ska finnas i dessa beskrivningar, vilka metadataelement av de valfria som bör ingå, huruvida ytterligare metadataelement bör läggas till beskrivningen och liknande hänsynstaganden. Inom arkivet och stadsförvaltningen har man ambitionen att förvaltningarnas material snabbt ska levereras till arkivet för att detta ska fungera som ett informationsnav för användarna inom staden. Eftersom vissa förvaltningar kan ha behov av att hålla kvar kopior av det levererade digitala materialet i sina verksamhetssystem kan det uppstå problem med versionshantering av de digitala objekten. Det är således viktigt att förvaltningarna inte ändrar något i de versioner av de digitala objekten de har kvar i sina system eftersom det skapar en autenticitetsproblematik kring objekten, därför har man avtalat att de får påbörja nya ärenden om de vill förändra något i de digitala objekten.

Leveranser till Riksarkivet görs inte i projektform på samma vis som hos de andra två institutionerna, istället underlättas arbetet för den levererande förvaltningen genom handledningar, framtagna metadatamallar och provleveranser via Riksarkivets leveransförberedelseverktyg (RALF). I likhet med överlämningar till de andra två arkivinstitutionerna ska levererande myndighet själva beskriva och dokumentera sina digitala objekt med metadata. I denna beskrivning ska ingå information om hur systemen var uppbyggda, vilka funktioner som fanns, vad olika poster och dokumenttyper innebar och dylikt. Informant 5 menar dock att metadatan som ska medfölja inte specificeras särskilt djupt, utan att man får anpassa sig till den metadata som finns tillgänglig i systemen. Informanterna menar att ju djupare systemet kan beskrivas desto bättre. Samtidigt exemplifierar informant 4 att de flesta myndigheterna använder samma ärendehanteringssystem och då behöver inte alla myndigheter beskriva hela detta utan endast de förändringar de gjort för att anpassa systemet efter deras behov. Hur djupt systemet kan beskrivas kan dock variera, det har hänt att vid leveranser från myndigheter som lagts ned att medföljande systemdokumentation endast kan bestå av systemmanualen och ingenting annat. Levererande myndighet ska konvertera de digitala objekten till godkända arkivformat innan leverans, det finns dock inga krav på att dokumentation av dessa migreringar ska medfölja leveransen. Loggar över hur de enskilda digitala objekten eller dokumenten hanterats i myndigheternas system överförs inte till arkivet då arkivet endast ska garantera autenticiteten för slutresultatet av dokumentationen, inte hur denna tagits fram menar informant 4. Detta kompletteras sedan med dokumentation om hur processerna i vilka de digitala objekten och dokumenten tagits fram och ärendehanteringssystemet i vilket dessa skapats.

Riksarkivet samarbetar medan uppsatsen skrivs med bland andra Sveriges kommuner och landsting och vissa statliga myndigheter i eARD (e-Arkiv och e-Diarium), ett

[58]Hemsida för e-arkiv Stockholm http://www.ssa.stockholm.se/Om-Stadsarkivet/E-arkiv-for-Stockholms-stad/ Tillgänglig: [2013-05-09].

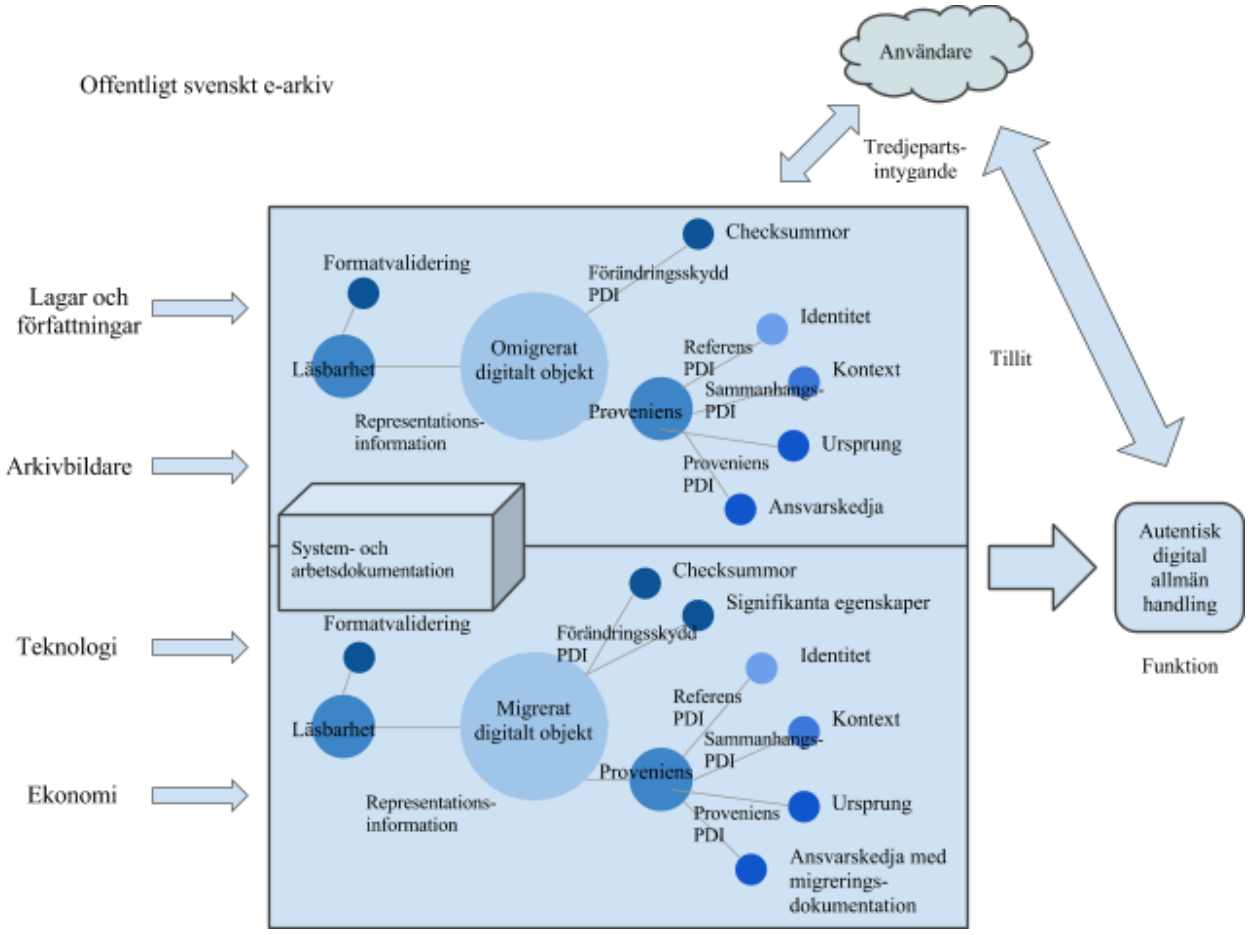
projekt som syftar till att ta fram rekommendationer för hur myndigheter kan anpassa sina informationssystem för att underlätta uttag och överföring till arkivinstitutioner. Genom projektet skapas bland annat förvaltningsgemensamma specifikationer (FGS:er) för exempelvis olika metadatautbytesformat, vilket ska göra det enklare att matcha metadatauttag från ärendehanteringssystem och liknande till metadatascheman inom arkiven.

9. Analys

Då den första delen av forskningsfrågan besvarades i resultatkapitlet syftar analyskapitlet till att besvara de två senare delarna av uppsatsens forskningsfråga *Vilka autenticitetsmarkörer för att bevara autenticitet hos digitala objekt används av offentliga arkivinstitutioner, varför har dessa metoder/metadata valts ut och vilka konsekvenser får detta urval för de framtida användarna?* Varför vissa autenticitetsmarkörer, metadata och metoder valts ut besvaras i avsnitt 9.2 och 9.3 medan de framtida användarnas perspektiv analyseras i 9.4. Kapitlets första avsnitt sammanfattar resultatet från det föregående kapitlet då det system som undersökningen identifierat presenteras.

9.1 Systemteoretisk återkoppling

Figur 17. Figur över undersökningsresultatet ur ett systemteoretiskt perspektiv.



Definitionen av det system som ämnades undersökas formulerades tidigare som: *Det offentliga e-arkivet är ett regelbaserat och formellt bevarandesystem med syftet att i ett långtidsperspektiv tillhandahålla funktionen läsbara och autentiska digitala allmänna handlingar.*

Systemet ovan har konstruerats efter resultaten av den empiriska undersökningen. Sammankopplingarna utgörs av de autenticitetsmarkörer som används av de undersökta institutionerna. Elementen i systemet är informationen i alla sammankopplingarna och de digitala objekten. Det är inte avgörande för systemet vilka allmänna handlingar som bevaras däri eller exakt hur exempelvis deras proveniens beskrivs, bara beskrivningen når upp till en autenticerande nivå. När tiden kommer för migreringar att genomföras kommer det att krävas nya sammankopplingar för att kunna leva upp till funktionen. Därför har systemet delats upp i två delar – en del med det omigrerade digitala objektet och en del med det migrerade digitala objektet. De nya sammankopplingar som kommer att krävas i det sistnämnda är signifikanta egenskaper och migreringsdokumentation. Vidare gör systemets avgränsning att det kan fungera autenticerande enligt tredjepartsintygandet. Då det är skilt från de som är "stakeholders" i dokumenten, exempelvis medborgaren och Skatteverket i deras korrespondens, kan denna neutralitet vara en garant för att dokumenten hanterats efter bästa förmåga. Systemteorin menar att systemet skapar något större än de enskilda delarna – tillsammans skapar de ett pålitligt system som tillgodoser behovet av autentiska digitala allmänna handlingar. Samtidigt skapar det faktum att ett etablerat system finns på plats tillit gentemot den information som bevaras där. Varje sammankoppling är dock viktig och bortfallet av endast en sådan får stora konsekvenser för funktionen – att bevara autentiska digitala allmänna handlingar.

Systemet är uppbyggt så att ny bevarandedokumentation, exempelvis tidsstämplar för olika åtgärder, kopplas till respektive autenticitetsmarkör. Till exempel faller information om att kontroller av checksummor gjorts och huruvida dessa är godkända under autenticitetsmarkören checksummor (eller i undantagsfall system- och arbetsdokumentation om institutionerna har de rutinerna – huvudsaken är att det finns plats i systemet för informationen). Ett alternativ vore istället att särskilja arkivets bevarandedokumentation i en egen markör, som exempelvis bevarandeloggar eller bevarandehistorik. Detta har dock nackdelen att det medför att de andra autenticitetsmarkörerna endast blir passiva ingångsvärden vilket minskar överskådligheten och användbarheten i systemet.

Det resultat som systemet ovan utgör är unikt i jämförelse med tidigare undersökningar gällande autenticitet.[59] En jämförelse mellan de autenticitetsmarkörer som ingår i detta system och de punkter som ett *record* måste uppfylla för att vara autentiskt enligt InterPARES 1-projektets resultat visar att de skiljer sig åt i flera hänseenden. I InterPARES checklista krävs exempelvis att fyra olika parter måste

[59] Såsom InterPARES 1 (2002c), Hänström (2007) och Bradley (2005) vilka beskrivits i kapitel 5.

vara namngivna[60] och att den aktivitet som dokumenteras i ett record måste namnges (InterPARES 2002a, ss. 5–8). Detta är inte aktuellt i denna undersöknings system av två anledningar. Den första kommer av att allmänna handlingar skiljer sig från records genom att de förra kan ha både ett inre och yttre bevisvärde medan de senare endast kan ha ett inre bevisvärde[61] (Hänström 2007, ss. 78–79). Detta medför att det inte är nödvändigt för systemet att ta hänsyn till ett krav på inre bevisvärde eftersom detta inte kommer att vara relevant för alla allmänna handlingar och då systemet beskriver en miniminivå som måste uppnås. Det inre bevisvärdet är dock inte helt utelämnat ur systemet eftersom det kan beskrivas under sammanhangs-PDI, men det är ingen egen autenticitetsmarkör vilket det hade varit i ett system som byggt på InterPARES resultat. Den andra anledningen till att skillnader uppstår mellan resultaten i InterPARES 1 och denna undersökning är att i den senare skiljs information som bevaras i en autenticitetsmarkör och autenticitetsmarkören i sig själv från varandra, ett tänkande som kommer av systemteorins användande av element och sammankopplingar. Detta teorival blir särskilt lyckat då man betänker att en åtskillnad mellan den information som bevaras och den institution som bevarar den kan härledas till den arkivvetenskapliga tanken att arkivet inte ska garantera att den information det bevarar är korrekt, bara att den är oförändrad (Hirtle 2000, s. 10; Lynch 2000, s. 36). Detta resonemang leder till att arkivet bara ska garantera att systemet finns implementerat och att det fungerar, men inte att den information som systemet fylls med är korrekt. Det sistnämnda ska istället garanteras av förvaltningarna. I InterPARES (2002a, s. 5) läggs ansvaret för autenticiteten istället på bevararen. Om det beror på andra juridiska förutsättningar, att de endast använt diplomatik för att tolka resultatet eller att de representerar en annan syn på arkivets funktion i samhället är osäkert.

Eftersom allmänna handlingar endast behöver ha ett yttre bevisvärde, vilket i sin tur gör att de inte ställs under samma specificerade krav som *records*, medför detta att det finns stora likheter mellan allmänna handlingar och digitala objekt så som dessa beskrivs av OAIS, eftersom dessa också alltid kan anses ha ett yttre bevisvärde och ibland också ett inre dito. Alla digitala objekt innehåller någon information som anses värd att bevara, även om denna information inte alltid dokumenterar någon särskild aktivitet. På så vis behöver uppsatsens system inte särskilja mellan digitala objekt som består i allmänna handlingar och digitala objekt som består av metadata som beskriver de allmänna handlingarna.[62]

Andra faktorer kopplade till det digitala bevarandet som har påverkat systemets uppbyggnad och utformning beskrivs i de följande avsnitten.

---

[60] Arkivbildare, författare, mottagare och "avsändare" (den sista skiljer sig från författare och arkivbildare så till vida att det är den som står som ägare till den programvara eller adress från vilket aktuell *record* skickats) (Interpares 2002a, s. 5).

[61] Det vill säga att de måste dokumentera någonting för att uppnå sin bevisstatus, till skillnad från det yttre bevisvärdet som inte har detta krav.

[62] Vilket måste anses vara en ganska oförutsedd konsekvens av tryckfrihetsförordningen (1949:105).

9.2 Bevarandestrategin påverkar systemet

Som konstaterats i resultatet har ingen av de studerade arkivinstitutionerna förlitat sig till varken emulering eller museimetoden som det sätt på vilket man vill tackla det digitala bevarandet. Enligt våra informanter, och även enligt litteraturen (van der Hoeven & van Wijngaarden 2005), är det främst i Nederländerna man har satsat på emulering som bevarandestrategi, men globalt sett verkar det vara ett ganska unikt ställningstagande. Museimetoden och emulering som bevarandestrategier har nackdelar i och med digital obsolescens eftersom den hård- och mjukvara som behövs för att strategierna ska vara funktionella blir utdaterade och även, gällande särskilt databärare och annan hårdvara, vittrar eller på andra sätt går sönder på grund av ålder. Dessutom kan det föreligga problem särskilt för museimetoden om man ser tillbaka till OAIS-definitionen av långtidsbevarande, nämligen att det ska vara oberoende förståeligt, alltså möjligt att förstå utan särskilda förkunskaper eller med hjälp av särskilda nyckelpersoner. Att använda museimetoden ställer dessutom höga krav på användarna eftersom de då måste ta sig till arkivinstitutionen och använda den gamla hårdvaran istället för att kunna ta del av de digitala objekten på andra platser.

Det är alltså inte förvånande eller anmärkningsvärt att de studerade institutionerna valt migreringsstrategin. Samtidigt kan man spekulera i huruvida det inte är just detta att man valt att använda sig av migreringsstrategin som gjort att *original* blivit ett mer meningslöst och tomt begrepp. Genom att anamma migreringsstrategin försätts institutionen i en position där man, i det långa loppet, inte längre kan hålla ett original[63] i form av en oförändrad bitström intakt eftersom man förbinder sig till att någon gång i framtiden migrera denna till ett nytt format. Performancemodellen ger att det är representationen på datorskärmen som utgör det som kan ses som originalet och att denna är beroende av hård- och mjukvara för att kunna uppfattas av en användare (Heslop et al 2002, ss. 8–10). Detta innebär att så länge representationen är densamma kan hård- och mjukvaran och även det digitala objektet varieras och bytas ut. Detta medför att det bästa mål man kan arbeta mot är istället att informationen ska bevaras, något som samtliga informanter menar.

En annan aspekt kopplat till information och arkivhandlingar är att i en digital kontext är inte informationen knuten till ett fysiskt medium på samma vis. Om ett digitalt objekt är lagrat på en fysisk databärare, exempelvis på en DVD-skiva, krävs det hård- och mjukvara för att kunna läsa den. Den är sålunda inte möjlig att tolka för en människa utan hjälpmedel. Detta är en stor förändring i jämförelse med pappersdokument som är omedelbart förståeliga. Då det fysiska mediet och databäraren blir mindre sammankopplat och därmed mindre viktigt kan man se att det blir viktigt att fråga "vad ska man bevara" för att kunna besvara "hur ska man bevara". Precis som Buckland (1998) menar så lönar det sig därför att se dokumentet mer funktionalistiskt i det digitala bevarandet, det som ska bevaras är snarare vad dokumentet gör eller är till för än själva det fysiska dokumentet. Ett dokument kan antas vara skapat för att förmedla eller föreviga någon sorts information, oavsett om det rör sig om något som är nedskrivet, fotograferat eller inspelat som ljud. Beroende

[63] Ur ett diplomatiskt perspektiv där originalet är den "förste i ordningen".

på vilket syfte den som ska ta del av och tolka informationen har kommer inte nödvändigtvis alla aspekter av dokumentet att vara menings- eller informationsbärande, utan vissa element kommer att vara viktigare än andra. För att dessa ska kunna bevaras digitalt måste de som står för det digitala bevarandet först och främst vara medvetna om att de finns för kunna bedöma deras betydelse. För att kunna bevara den fysiska strukturen och andra visuella informationsbärande element kan det bli nödvändigt att beskriva och bevara dessa som signifikanta egenskaper, det kan även gälla intrinsic element för att anknyta till diplomatiken. Hade man istället valt emuleringsstrategin hade look and feel automatiskt varit intakt eftersom att representationen ska vara exakt, eller så lik som möjligt, den ursprungliga är en bärande tanke bakom emuleringskonceptet. I denna strategi är också originalbitströmmen säkrad eftersom den aldrig genomgår några förändringar, arbetet med det digitala bevarandet syftar istället till att skapa programvara som kan läsa och återge det digitala objektet, snarare än att anpassa det digitala objektet till programvaran.

9.3 Varför har vissa autenticitetsmarkörer valts ut?

Vissa markörer är mer utpräglat tekniska och har därför lagts till de metoder som användes även inom analogt bevarande, även om de förvisso inte självklart explicit sågs som kopplade till autenticiteten hos arkivhandlingarna. De tekniska markörerna redovisas först i avsnittet.

Då migreringsstrategin valts skapas vissa förutsättningar för vilka autenticitetsmarkörer som blir aktuella att använda. Samtliga arkivinstitutioner använder checksummor på olika nivåer för att kontrollera och garantera att integriteten hos det digitala objektet är oförändrat. Detta ökar sannolikheten för att informationens integritet också ska vara intakt, men det förutsätter att man använder mjukvara som kan läsa och återge det digitala objektet så som det är tänkt. Även om institutionerna bevarat enligt museimetoden eller emuleringsstrategin hade checksummor varit aktuella att använda eftersom det är den enda metodik som garanterar att ett digitalt objekt är oförändrat ner på bitnivå.

Om man ser ett digitalt objekt som en burk hallonsylt dokumenterar och kontrollerar checksumman att varje frö ligger kvar, att det är samma burk, med samma lock och att etiketten är densamma, medan en formatvalidering kontrollerar vad som står på etiketten och ser om burken väger så mycket som den ska. De två metoderna är således något besläktade med varandra såtillvida att de båda kontrollerar integriteten hos det digitala objektet, även om checksumman gör det mycket mer noggrannt. De undersökta institutionerna använder sig inte av formatvalidering primärt i sitt digitala bevarande eftersom de, i de flesta fall, antingen kan säga till förvaltningarna att de vill ha särskilda burkar,[64] det vill säga filformat, eller att de själva konverterar innehållet.[65] Dock sker en formatvalidering när de digitala objekten ska läggas in i e-

[64] Vilket Riksarkivet och Stockholms stadsarkiv gör.
[65] Vilket SYLL gör.

arkivplattformen eftersom denna endast godkänner vissa filformat enligt de minsta gemensamma nämnare som valts ut. Detta kan vara fördelaktigt i ett långt tidsperspektiv eftersom det är osäkert om dessa små förändringar i filformatsuppbyggnaden skulle kunna få konsekvenser vid framtida migreringar eller att viss mjukvara inte skulle kunna tolka de digitala objekten i en avlägsen framtid. Tillsammans lägger dessa två markörer som främst är kopplade till integriteten hos de digitala objekten grunden för att arkivet ska ha en dokumenterad startpunkt för det digitala objektets integritet vid leverans till arkivet, eller efter varje migrering. Formatvalideringen är också kopplad till den framtida läsbarheten hos det digitala objektet, vilket är en förutsättning för att informationen ska vara bevarad och autentisk. I fallet med de studerade institutionerna dokumenteras formatvalideringen inte, men den är en förutsättning för att de digitala objekten ska kunna läggas till i arkivet vilket gör att den dokumenteras implicit med att objekten inkorporeras i arkivet, men inte explicit i metadata. Den metadata och funktionalitet som formatvalideringen kontrollerar ingår i representationsinformationens undergrupp semantisk och övrig metadata enligt OAIS-klassifikationen.

Ingen av de tre studerade arkivinstitutionerna uppger att de använder digitala signaturer. Vad som utgör en digital signatur verkar inte helt enhetligt, varken i litteraturen (McNeil 1998, s. 112; Boudrez 2007, ss. 179ff; Gladney 2007, ss. 167ff; Adam 2010, ss. 601–602) eller bland informanterna, och de olika synsätten eller varianterna får olika konsekvenser:

En digital signatur som innebär att ett digitalt objekt krypteras för alla utom den som har krypteringsnyckeln, vilket alltså inte skulle vara arkivet, medför att det kan uppkomma problem vid långtidsbevarandet eftersom man kan fråga sig hur man kan kontrollera huruvida en migrering varit framgångsrik om man inte kan läsa objektet då det är krypterat. Det är även oklart om man kan migrera ett krypterat material, eller om krypteringen gör att migreringen korrumperar det digitala objektet.

Om den digitala signaturen däremot består i ett okrypterat digitalt objekt med en krypterad checksumma uppstår ändå problem när det digitala objektet migreras, eftersom checksumman då inte längre är giltig. Detta gör i så fall att den digitala signaturen som ett diplomatiskt sigill och garant för att integriteten hos det digitala objektet är oförändrad blir meningslös eftersom det de facto har skett en integritetsomvälvande händelse i och med migreringen. Digitala signaturer som underskrifter, exempelvis som autenticering av bankärenden via en internetbank med hjälp av en e-legitimation, är också problematiska. Detta eftersom det blir mycket som måste bevaras kring det system som skapade dem och länkar mellan detta och det signerade digitala objektet för att de ska kunna autenticeras. Om denna information kan inkorporeras i metadatan kring det digitala objektet på ett autentiskt riktigt sätt, till exempel genom en logg eller metadata kopplat direkt till objektet hade det inte blivit lika problematiskt.

Oavsett vad man ser som digitala signaturer medför de fler problem än vinster för det digitala bevarandet menar informanterna, särskilt i ett långtidsperspektiv, vilket har gjort att de studerade arkivinstitutionerna valt att inte använda dem. Dock kan man tänka sig att eftersom inklusionskriterierna specificerade att det studerade arkivet skulle vara ett mellan- eller slutarkiv, och samtliga tre snarare är slutarkiv, kan detta

inverka på varför samtliga studerade institutioner valt bort digitala signaturer. Inom förvaltningarna och i vissa typer av mellanarkiv är det möjligt att de skulle vara mindre problematiska eftersom tidsperspektivet är kortare, vilket bör medföra färre problem med digital obsolescens och kopplingar till andra system.

På en metadatanivå ingår alla de tre tekniska autenticitetsmarkörerna i den metadata som inom OAIS-modellen kallades PDI, preservation description information, närmare bestämt i undergruppen förändringsskydd. Kopplat till de tekniska markörerna är dokumentationen av dessa i exempelvis bevarandemetadataschema eller andra typer av loggar över vilka bevarandeåtgärder som vidtagits. Den typ av bevarandehistoriksmetadata som då uppkommer ingår i PDI-undergruppen *proveniensinformation.* Särskilt de tekniska autenticitetsmarkörerna och hur de kontrolleras i arkivet vid olika tillfällen kan genereras med hjälp av *automatisk* metadata eftersom de på en grundläggande nivå[66] inte måste motiveras eller förklaras, vilket hade krävt mänsklig inblandning. Denna dokumentation över vilka bevarandeåtgärder som vidtas angående de digitala objekten är viktig eftersom den är det enda som visar att något gjorts och hur det gjordes. Förutom automatisk metadata kan vissa åtgärder behöva förklaras utförligare, exempelvis då en formatmigrering måste göras och en utredning vidtas över vilka förändringar denna medför för de digitala objekten. Dokumentationen över utredningen behöver inte nödvändigtvis inkorporeras i varje bevarandemetadatafil, men bör på något vis kopplas till de digitala objekten på ett sätt som står sig över långa tidsrymder. Om denna koppling är till varje enskilt digitalt objekt eller till hela beståndet kan variera beroende på förutsättningarna, både gällande detaljrikedom i dokumentationen, vilka datamängder som uppstår, hur säkra kopplingarna eller länkarna mellan de digitala objekten och dokumentationen kan göras, et cetera.

Som tidigare nämnts har inga av de tre studerade institutionerna ännu ställts inför arkivformatmigreringar och har heller inte formaliserat några rutiner inför detta eftersom de menar att det är bättre att ta ställning till det när man vet vilka förutsättningarna blir. Man ser dock liknande lösningar på problemet med hur informationen ska bevaras intakt över en, eller flera, formatmigreringar. Vid leveranserna till arkivet ska levererande myndighet skicka med dokumentation över hur informationen presenterades och såg ut i de ursprungliga systemen, och denna dokumentation kan ligga till grund för en bedömning om vilka element som är mer informationsbärande än andra. Dessa kan sedan dokumenteras, antingen direkt som signifikanta egenskaper i ett metadataschema eller i den obearbetade formen som systemdokumentation och skärmdumpar, så länge de tas i beaktande vid formatmigreringen. Signifikanta egenskaper och de bevarandeåtgärder som associeras med dem hör till PDI-undergruppen proveniensinformation.

Arkivinstitutioner har sedan många år utvecklat en *best practice* kring hur man ska bevara autenticiteten hos analogt arkivmaterial. Autenticitetsmarkörerna för ett analogt material består i hur proveniensen kan visas och hur utförligt den beskrivs,

[66] Sker en åtgärd eller inte, i så fall vid vilken tidpunkt, vilket program eller algoritm används och så vidare.

dels gällande vem som är skapare till arkivhandlingen, men också varför den skapats och vem som haft ansvar för den från skapelsen tills att den begärs ut. Som PDI beskrivs proveniensen av undergrupperna proveniens-, referens- och sammanhangsinformation. Referensinformationen täcker in de aspekter som det digitala objektets namn, beteckning, diarienummer, och så vidare. Sammanhangsinformationsmetadatan beskriver hur det digitala objektet relaterar till andra digitala objekt och strukturer inom arkivet, till exempel med hjälp av arkivförteckningar eller arkivbeskrivningar. Tillsammans med referensinformationen dokumenterar denna kontexten kring det digitala objektet och dess information, vilket i sin tur genererar ytterligare information kring det digitala objektet. För att återknyta till Hirtle (2000, s. 10) är det just denna kontextuella koppling som är det fundamentala för arkivens roll i bevarandet. Ansvarskedjan, det vill säga vem eller vilken institution som haft ansvaret för de digitala objekten faller istället under proveniensinformation. Det är ur ett autenticitetperspektiv viktigt att alla dessa tre PDI-undergrupper används för att beskriva det digitala objektet eftersom det annars blir omöjligt att fullt ut kunna påvisa dess autenticitet ur ett diplomatiskt perspektiv eftersom detta kräver att dokumentet ska kunna identifieras och sättas i sitt sammanhang. Att arkivet kan redovisa proveniensen hos digitala arkivhandlingar enligt dessa principer är en förutsättning för att institutionen ska kunna anses leva upp till *best practice* inom området, vilket är ett steg i rätt riktning om man ser till Moss (2008, s. 80) krav att arkivet måste visa att det för sig *beyond reproach* för att förtjäna tillit. *Best practice* och tillit är kopplat till tredjepartsintygandet, det vill säga det autenticitetsvärde som kan associeras till en arkivhandling på grund av att den bevaras av ett arkiv. Tredjepartsintygandet blir således inte en egen autenticitetsmarkör, men den ligger implicit kopplad till ansvarskedjan och den dokumenterade redovisningen av bevarandet, både genom bevarandemetadataschema eller loggar, och dokumentation av institutionens rutiner och arbetsmetoder. I de fall dessa lever upp till de krav som kan ställas angående transparens och *best practice* ökar trovärdighet och tilliten för institutionen, vilket gör att det kan ses som meriterande för det digitala objektet att det bevarats i dess arkiv. Informant 1 poängterade vikten av att ett arkiv måste arbeta mer proaktivt med det digitala materialet och bevarandeplaneringsfunktionen i OAIS-modellen för att kunna arbeta enligt *best practice*. Det blir svårt för ett digitalt bevarandesystem att negligera ny teknik och nya format under en längre tid och ändå behålla trovärdigheten gällande bevarandet.

Tidigare forskningsprojekt såsom InterPARES 1 visade, vilket sedan bekräftades i intervjuerna, att en viktig brytpunkt för autenticiteten hos ett digitalt objekt uppkommer vid överlämnandet till mellan- eller slutarkiv från det system i vilket det skapats. Det framkom vidare i intervjuerna att det inte är ovanligt att digitala objekt migreras mellan olika informations- och ärendehanteringssystem flera gånger under sin aktiva period ute på myndigheter och förvaltningar. Dessa migreringar är inte alltid så kontrollerade att de tar hänsyn till signifikanta egenskaper och dylikt, och de dokumenteras ofta inte så utförligt. I alla fall inte på ett sätt som sedan kommer arkivinstitutionen och dess användare till godo eftersom det är ovanligt att bevarandemetadataschema som egna digitala objekt medföljer i arkivleveranser till de studerade arkivinstitutionerna. Enligt Hänström (2007, s. 83) tar den autenticitetsbevarande dokumentationen sin början vid leverans till arkivet eftersom det är där denna typ av metadata sätts till de digitala arkivhandlingarna. För att få

ökad kontinuitet i det digitala bevarandet håller eARD-projektet nu på att ta fram förvaltningsgemensamma specifikationer som ska underlätta utbytet av metadata kring digitala objekt som levereras till arkiven. Genom att underlätta detta kan förhoppningsvis kraven i FGS:erna successivt höjas till en nivå där PDI som är likvärdig den som sätts till de digitala objekten av arkivet tillkommer redan på förvaltningarna.

Oavsett om detta realiseras i framtiden kvarstår det att de val som görs angående vad som ska dokumenteras och bevaras kopplat till de digitala objektens autenticitet får konsekvenser för de framtida användarna.

9.4 Institutionernas val och de framtida användarna

Eftersom alla de studerade arkivinstitutionerna är offentliga och innehåller allmänna handlingar menade ingen av respondenterna att de kunde rikta in sig mot en särskild målgrupp, vilket annars förordas av OAIS-modellen där långtidsbevarande definieras som beroende av en *designated community* (CCSDS 2012, 1-11). Denna definition av bevarande blir delvis inkongruent med traditionellt arkivvetenskapligt tänkande om det målgruppsinriktade bevarandet skulle gå så långt att arkivhandlingarna förtecknas eller sorteras på ett sätt som bryter ut dem ur sin originalkontext. Samtidigt kan det vara viktigt att hålla i minnet att OAIS-modellen inte främst skapats för att bevara *records* eller allmänna handlingar, utan enorma mängder forskningsrådata, ofta i databaser och tabeller.

Metadatan som är kopplad till de digitala objekten i de studerade arkivinstitutionerna kan dock vara specialiserad så att vissa användargrupper har fördelar jämfört med andra för att förstå de digitala objekten. Exempelvis planeras att forskningsdatan som ska bevaras inom SYLL ska sättas till i samarbete med de forskare som skapat leveransen, vilket i förlängningen gör att det kommer att vara anpassat efter det forskningsområde i vilket det uppkommit – något som egentligen ter sig ganska naturligt. Dock medför detta att det kan vara svårare att återanvända materialet inom andra forskningsdiscipliner. Arkivinstitutionerna ser istället fördelar med att ta emot så mycket metadata och dokumentation förknippade med de digitala objekten som går att leverera, förutom loggar.

Loggar över hur digitala objekt hanterats inom till exempel ett ärendehanteringssystem är intressanta ur ett arkivvetenskapligt hänseende. Enligt Hirtle (2000, s. 10) är det bara arkivets uppgift att garantera autenticiteten hos arkivmaterial från och med att det inkommit till arkivet, det är alltså bara det dokumenterade *slutresultatet* av förvaltningarnas aktiviteter som ska bevaras autentiskt- inte de arbetsprocesser och händelser som dessa dokumenterar. Detta är ett bra argument för att offentliga arkivinstitutioner inte ska ta emot loggar över hur de allmänna handlingarna skapats och hanterats på myndigheterna, och skiljer sig heller inte från hur man bevarar analogt material. Dock kan man ställa sig frågan om vad utgångspunkten för Hirtles resonemang är: kan det vara att i ett analogt perspektiv är det svårt att dokumentera hur handlingarna uppkommit och därför har beslutet fattats att denna dokumentation, som i de fall den finns ofta klassas som arbetsmaterial och gallras, inte ska medfölja till arkivet? Eller är det ett utslag av att autenticiteten hos resultatet, den allmänna handlingen, anses garanterad genom ett slags

tredjepartsintygande som kommer av att den är upprättad av en viss myndighet som det finns förtroende för? Det senare alternativet blir så klart problematiskt i de fall det visar sig att en myndighet misskött sitt arkiv eller sina dokumentations- eller ärendehanteringsrutiner.

En annan anledning som angavs till att loggar inte bevaras är att de inte alltid är funktionella; för att en logg ska vara intressant att bevara ur autenticitetssynpunkt är det viktigt att den endast loggar händelser av relevans, eftersom det annars lagras mycket "brus" vilket gör att den blir svår att navigera i. Samtidigt måste detta urval av vilka händelser som loggas göras med tanke på ett framtida användningsområde, vilket implicit innebär att man måste se till en potentiell användare och dess behov. Det blir också lite paradoxalt om loggar inte tas emot eftersom de innehåller för mycket information, medan annan metadata inte avböjs med motiveringen att framtida användare får sålla fram det de själva behöver. Det kan därför vara sannolikt att det är främst kopplingen till hur det sett ut i analogt bevarande gällande arbetsmaterial och dylikt som gör att arkiven inte tar emot loggar.

Vilken anledningen än må vara så får den konsekvenser för den framtida användaren eftersom det gör att denne inte kan spåra autenticiteten hos ett digitalt dokument tiden innan det inkom till arkivet, även om denna information skulle kunna bevaras. Ur ett diplomatiskt och funktionalistiskt perspektiv kan man återigen fråga sig vad det är som ska bevaras? Ser man själva dokumentet eller den allmänna handlingen som central så är det detta slutresultat som är det primära att bevara, frågan är om det blir lika självklart om man ser till att det är själva informationen som ska bevaras. Ur ett hermeneutiskt perspektiv, där hur världen uppfattas beror på förförståelse, kan det vara informativt och värdefullt att bevara loggar eftersom de visar hur informationen har kommit till och på så vis kan ge ledtrådar om författarens värderingshorisont,[67] vilket kan utgöra ett intressant forskningsområde i sig.

Även om loggar från myndigheternas ärendehanteringssystem inte sparas är loggandet av bevarandeåtgärder och händelser desto viktigare inom e-arkivet för informanterna och deras institutioner. Allt som görs med de digitala objekten, från att en checksumma tillsätts eller kontrolleras till migreringar, dokumenteras. Detta medför att, så länge dessa bevarandemetadata- eller arbetsloggar verkligen täcker in alla viktiga händelser, har den framtida användaren möjlighet att se exakt hur ett digitalt objekt handhavts av arkivet vilket ökar dess autenticitet enligt principen för tredjepartsintygande.

Att en checksumma sätts till ett digitalt objekt och att denna sedan bevaras för framtiden tillsammans med dokumenterade kontroller över huruvida den förändrats görs av samtliga studerade arkivinstitutioner. Detta möjliggör för användaren att kontrollera om integriteten hos det digitala objektet är intakt mellan migreringarna. Vid en migrering slutar den gamla checksumman att gälla och en ny måste beräknas.

---

[67]Ett centralt begrepp inom hermeneutiken som innebär, för att citera Fredriksson (2003, s. 38) *"*mängden av de medvetna och omedvetna uppfattningar och hållningar som vi hyser vid en given tidpunkt, och som vi inte riktar vår uppmärksamhet mot".

För att användaren ska kunna förvissa sig om att integriteten hos det digitala objektet är oförändrat efter en migrering måste de förändringar som denna givit upphov till redovisas och dokumenteras på något sätt. Även om informationsbärande element för den framtida användaren fallit bort och inte bevarats som signifikanta egenskaper är det viktigt för transparensen och autenticitetsbedömandet att användaren kan ta del av hur den ursprungliga representationen såg ut.[68] Hur signifikanta egenskaper kommer att användas och dokumenteras vid framtida migreringar är inget som institutionerna tagit ställning till i dagsläget, även om det finns möjlighet och utrymme för det enligt de bevarandemetadatascheman som används. Med den teoribildning, metodik och teknologi som finns tillgänglig idag ter det sig dock som att utredningar kring vad som är signifikanta egenskaper i olika handlingsslag och bevarandet av dessa är en förutsättning för att integriteten hos ett digitalt objekt ska kunna bevisas för en framtida användare. Kanske blir det i framtiden möjligt för nya datorprogram att komma runt problemet med att det endast är genom att människor gör visuella kontroller som man kan kontrollera huruvida signifikanta egenskaper bevarats. Eftersom den förutsättningen dock inte finns i nuläget ter det sig som om att man aldrig med 100 % säkerhet kan säga att inga betydelsefulla informationsförluster kommer att uppkomma i ett stort arkivbestånd som migreras eftersom stickprover aldrig kan garantera hela beståndet.

För att framtida användare ska kunna kontrollera proveniensen hos ett digitalt objekt måste denna vara noggrannt beskriven. Dels måste sådan referensinformationen finnas med för att klart kunna identifiera det enskilda objektet, men sedan behövs sammanhangsinformationen för att sätta in det i sin kontext inom arkivet, men även att sätta arkivet och arkivbildaren i kontext gentemot omgivningen. Som nämndes i bakgrundskapitlet blir förtecknandet av arkiv med nödvändighet subjektivt och arkivariens verklighetsuppfattning och förståelse färgar arbetet (Duff & Harris 2002, ss. 280ff). Arkivbeskrivningar är produkter av sin tid vilket gör att över ett långt tidsspann kan deras informationsinnehåll förändras avsevärt. Enligt den infologiska ekvationen så kommer de kunskaper och den förförståelse (S) som användarna besitter och använder i sin tolkningsprocess att ha förändrats som ett resultat av tidsvariablen (t).[69] Diplomatiken identifierar de viktigaste proveniensattributen som identitet/ursprung och kontext vilket kan ange en lägsta nivå som måste uppnås för att proveniensen ska anses redovisad. Samtidigt följer av Duff och Harris' resonemang och den infologiska ekvationen att kontext och identifikation kan vara subjektivt och beroende av användares förförståelse. Detta har traditionellt sett inneburit att de analoga arkiven har lagt över en hel del av ansvaret att förstå arkivmaterialet på användarna, även om hjälpmedel tagits fram för att underlätta. I en digital kontext, med OAIS-modellens definition av långtidsbevarande blir detta dock svårare. Främst eftersom denna innebär att den bevarade informationen ska vara oberoende förståelig medför detta otroligt höga krav på detaljnivån i arkivbeskrivningarna om dessa inte riktar sig till en avgränsad målgrupp, vilket ju inte är möjligt att göra för de studerade

[68] Exempelvis via skärmdumpar som ju ingår i det material som ska ingå i leveranserna till arkiven enligt informanterna.
[69] Vilken i sin tur medför att användarna inte kommer att vara desamma om tidsspannet överstiger en mänsklig livslängd.

arkivinstitutionerna då de bevarar allmänna handlingar vars målgrupp är allmänheten, forskningen och förvaltningarna. Det blir också svårare i de fall e-arkivets ambition är att vara ett informationsnav som är användarvänligt och inbjudande, likt Stockholms stadsarkiv.

Viktigt att hålla i minnet gällande deskriptiv information och metadata kopplad till digitala arkiv är att kopplingarna eller länkarna, både mellan digitala objekt och olika hierarkier inom arkivet och externt till exempelvis webbplatser, också måste bevaras och autenticeras. I de fall dessa kopplingar till andra digitala eller analoga informationsresurser skulle korrumperas räknas detta också som en informationsförlust som även påverkar autenticiteten såtillvida att kontextuell information som kan vara kopplad till proveniensen försvinner. Om all denna kontextuella information försvinner, eller förminskas, gäller detta även den autenticitet som förlänas det digitala objektet genom tredjepartsintygandet. Detta eftersom arkivets bevaranderutiner kan ifrågasättas, men också för att det är den kontextuella kopplingen som skapar en arkivhandling enligt Hirtle (2000, s. 10). Metadatan som utgör kopplingarna och länkarna till andra digitala resurser eller digitala objekt kan i sig själva anses utgöra digitala objekt som i sin tur behöver autenticerande metadata. Denna metadata kan sedan behöva autenticerande metadata, vilket kan fortgå i en oändlig slinga. Denna slinga kan teoretiskt sett brytas först när användaren menar att den inte längre måste fortgå för att de ska lita på autenticiteten hos det digitala objektet. I realiteten bryts den oftast efter högst en iteration; checksummor kan förutom att vara kopplade till ett digitalt objekt sättas till ett helt informationspaket, sparas på flera fysiska platser[70] eller att gamla checksummor bevaras trots att nya satts till, men det verkar vara sällsynt att metadatan för en checksumma förses med en checksumma och så vidare. I litteraturöversikten presenterades flera olika typer av migreringar, varav den som kallades transformeringar, det vill säga formatmigreringar hos det digitala objektet, är det som behandlats mest i uppsatsen eftersom det är den som har störst påverkan på integriteten hos det digitala objektet. Transformeringar kan även inbegripa att det digitala objektets metadata migreras till ett annat format samtidigt som det digitala objektet självt migreras, men det finns även en migrationsform, ompaketering, som bara innebär en migrering av själva metadatan. Denna kan medföra den typ av kontextuella informationsförluster som beskrivs ovan. Ett annat problem som kan uppkomma över långa tidsrymder är att externa länkar till webbplatser inte fungerar eftersom dessa kan ha bytt adress, eller att den resurs de länkar till lagts ned. I de fall den kontextuella proveniensinformationen är väldigt beroende av externa länkar kan det få konsekvenser för autenticiteten. Vad händer till exempel om en extern länk pekar mot en resurs som byts ut mot en version som inte är kompatibelt med det digitala objektet eller om resursen byts ut mot en förfalskning? Arkiven kan ju trots allt inte garantera autenticiteten hos annat än det material de själva bevarar.

[70] Detta sätt att omsätta LOCKSS-metodiken, alltså att spara flera kopior på olika platser för att kunna jämföra dem, gällande checksummor används utav Riksarkivet för att öka säkerheten, kunna kontrollera om checksummor korrumperats och bevisa deras autenticitet.

Proveniens- och integritetsaspekten ansvarskedjan kan bevaras antingen genom att den dokumenteras deskriptivt i arkivbeskrivningar eller automatiskt i bevarandeloggar.[71] Detta kan göras mer eller mindre utförligt; från att endast identifiera vilken institution som haft ansvar för att beskriva och bevara vilka åtgärder de vidtagit gällande arkivmaterialet. Detta är inte endast upp till den arkivinstitution som slutligen bevarar arkivmaterialet att avgöra eftersom det beror på vilka bevaranderutiner de som tidigare haft ansvar för arkivmaterialet haft och vilken dokumentation de kan leverera vidare. Eftersom ingen av de studerade arkivinstitutionerna kan kräva att de levererande myndigheterna, förvaltningarna eller arkivbildarna bifogar den här typen av dokumentation finns det en möjlighet att de framtida användarna inte kommer att ha tillgång till detaljerad information om ansvarskedjan genom hela arkivhandlingarnas livscykel. Istället får de lita till förhoppningsvis utförliga arkivbeskrivningar och andra källor utanför arkivpaketet för att få mer information om de som tidigare ansvarat för arkivhandlingarna och vilka åtgärder de vidtagit i bevarandet. Hur detaljerad information om de tidigare förvaltarna användaren uppfattar sig behöva har återigen med tillit att göra, litar användaren tillräckligt på den tidigare förvaltaren för att vara säker på att arkivhandlingen inte korrumperats innan den inkommit till arkivet?

Sammanfattningsvis använder de studerade arkivinstitutionerna samma autenticitetsmarkörer, även om de inte grupperas eller kopplas till det digitala objektet på samma sätt. Dessa täcker in alla de metadataundergrupper som beskrivs av OAIS-modellen vilket medför att de digitala objekten beskrivs utifrån flera kompletterande aspekter gällande autenticitet, integritet och proveniens.

[71] Bevarandeloggar och de loggar som gallras skiljer sig åt såtillvida att loggarna dokumenterar förändringar *användarna* gör i ärendena i ärendehanteringssystemet, medan bevarandeloggarna dokumenterar bevarandeåtgärder som vidtas gällande både ärendena och ärendehanteringssystemet i sin helhet och görs av IT- eller informationsansvarig personal.

10. Diskussion

Diskussionskapitlet är delat i två avsnitt: i det första diskuteras undersökningens resultat i förhållande till forskningsfrågan, det teoretiska perspektivet och omvärlden. Det andra avsnittet diskuterar metodvalet i förhållande till undersökningen och dess resultat.

10.1 Resultatdiskussion

Hade de två förfalskningsexemplen från inledningen som användes för att belysa hur viktigt det är att kunna kontrollera om en arkivhandling är autentisk kunnat inträffa i en digital kontext? Uppsatsen har inte behandlat informationssäkerhet såtillvida att arkivens rutiner för IT-säkerhet och -attacker undersökts,[72] så frågan om hur stor risken är för att ett e-arkiv ska kunna korrumperas av parter utanför arkivet kan inte besvaras fullt ut. Dock kan man ur ett diplomatiskt och OAIS-perspektiv se att, till skillnad från i det analoga arkivet, handhar användaren aldrig originalet. Dels eftersom det är svårt att säga vad som är originalet i det digitala, dels för att även om en DIP lämnar ett arkiv så finns kopior av de digitala objekt kvar i en eller flera AIP i e-arkivet. Alltså kan inte en användare lägga till fler handlingar som i det första exemplet,[73] eller ändra en fyra till en femma som i det andra. En större risk för att digitala arkivhandlingar eller digitala objekt ska förlora sin autenticitet är istället kopplat till arkivets långtidsbevarande och rykte. För att minimera risken med denna sorts autenticitetsförlust är det viktigt att arkivet dokumenterar de digitala bevarandeåtgärder som vidtas och kopplar dessa till de digitala objekten i bevarandemetadata, arbetsloggar eller annan dokumentation.

Transparensen gentemot den framtida användaren är en förutsättning för att denna ska känna tillit till arkivinstitutionen. Detta i sin tur är en förutsättning för att bryta de potentiellt ändlösa autenticitetsiterationerna med dokumentation som beskriver hur dokumentationen om hur dokumentationen uppkommit och så vidare, som annars krävs för att autenticera arkivhandlingarna. Tillit är även en förutsättning för det autenticerande tredjepartsintygandet och att autenticitet bevaras genom ansvarskedjan. Det är även av vikt att arkivinstitutionen arbetar proaktivt med e-

[72] Dock är Riksarkivets bevaranderutiner att spara kopior av checksummor och digitala objekt på flera olika platser i landet ett sätt att arbeta med IT-säkerhet. Likaså det faktum att innehållet i e-arkivet ligger på skyddade servrar som inte är kopplade till internet och att användarna inte har fri tillgång till detta utan leveranser måste skötas via en mottagandefunktion och utlämnande genom en åtkomstfunktion enligt OAIS-modellen.
[73] Såvida inte den som levererar inkluderar förfalskningar till arkivet dolda bland mängder av andra arkivhandlingar.

arkivet och håller sig uppdaterade enligt bevarandeplaneringsfunktionen i OAIS för att kunna bevara enligt samtida *best practice* och kunna bemöta svårigheter i bevarandet, såsom teknisk obsolescens och nya standardarkivformat. Även på en mer detaljerad nivå i bevarandet av de digitala objekten är det viktigt att e-arkiven är proaktiva: trots att inga större formatmigreringar företagits på någondera av de undersökta institutionerna och inga rutiner kring detta formaliserats var informanterna överens om att inför migreringar måste undersökningar kring dessas effekter göras och efter migreringen måste resultatet kontrolleras. Formatmigreringar kan förvanska den logiska och fysiska strukturen, eller det som diplomatiken kallar *intrinsic* och *extrinsic elements*, vilket påverkar autenticiteten hos informationsinnehållet – det vill säga integriteten hos det digitala objektet. Eftersom det centrala att bevara sågs som informationen innebär detta att bortfall av informationsbärande element i de digitala handlingarna är en typ av ofrivillig gallring (RA-FS 2009:2 2 kap. 1 §). Samtidigt kan migreringar av metadatan göra att kontextuell information, till exempel interna eller externa länkar, kopplad till det digitala objektets proveniens försvinner, vilket också är ofrivillig gallring. Kan inte autenticiteten bevisas räknas det digitala arkivmaterialet också som ofrivilligt gallrat, vilket sätter autenticitetens vikt i det digitala bevarandet i fokus.

De digitala objekt som skapas idag och som plockas fram av användare i nästa århundrade kommer med största sannolikhet ha migrerats åtminstone en gång. Vissa av de autenticitetsmarkörer som dessa kommer att ha tillgång till befinner sig på en graderad skala, medan andra är mer absoluta. Till de sistnämnda hör checksummorna, antingen stämmer de överens eller inte. De markörer som befinner sig på en graderad skala är desto fler, till dessa hör alla de deskriptiva markörerna såsom arkivbeskrivningar eftersom dessa kan vara mer eller mindre informationsrika och att denna information kan vara mer eller användbar för den framtida målgruppen. Om integriteten är helt intakt och det digitala objektets informationsinnehåll inte ändrats har inget gått förlorat. Dock är det troligt att man vid migreringar får välja att bara behålla vissa signifikanta egenskaper, medan andra karakteristika i dokumentet gallras. Bortfallet av vissa karakteristika behöver inte innebära att det digitala objektet är inautentiskt, bara dessa gallrade egenskaper dokumenteras och att de har existerat redovisas.[74] Man kan då inte argumentera för att det digitala objektet inte *är vad det utger sig för att vara*, det autenticitetsmaxim som flertalet autenticitetsdefinitioner faller tillbaka på. Dock härleds denna autenticitet från att det digitala objektet utger sig för att vara ett *annat* digitalt objekt än det ursprungliga. De förändringar som uppkommit gör att det bör ses som en sorts *imiterande kopia*[75] av dess förlaga, snarare än identisk eller likvärdig denna. Detta gör att det ter sig rimligt att dela undersökningens identifierade system[76] i två halvor där den ena halvan innehåller migrerade och den andra omigrerade digitala objekt. Även om de omigrerade digitala objekten i någon mening också är kopior så är de *kopior i samma form som*

[74] Vilket påminner om gallring av arkivhandlingar där gallringsutredningar och de gallrade handlingarnas diarieposter bevaras som ett bevis på att handlingarna existerat och ger insyn och legitimitet åt att de gallrats.
[75] Exempel på imiterande kopior är enfärgade kopior av flerfärgsoriginal. Se avsnitt 7.4.2, integritet ur ett diplomatiskt teoretiskt perspektiv.
[76] Se avsnitt 9.1.

originalet, det vill säga att likt kontrakt upprättade i flera exemplar är tiden för deras skapelse det enda som skiljer dem åt. Eftersom förutsättningarna för migrerade och omigrerade digitala objekt enligt ovanstående resonemang skiljer sig åt till så hög grad skiljer sig också möjligheterna för bedömningen av deras autenticitet åt. Till syvende och sist är all autenticitet beroende av bevis och även om autenticiteten kan vara absolut[77] på ett filosofiskt plan är det bevisen och *deras autenticitet* som bidrar till att autenticitet är ett graderat värde.

Beroende på vilken syn man har på autenticitet hos digitala dokument kommer de metoder, tekniker eller arbetssätt som kan bevara autenticitet att vara garanter för denna i varierande grad. Ett synsätt där man lägger tonvikten vid tredjepartsbevarandet kommer att leda till att den viktigaste aspekten är att visa och bevara tilliten för institutionen som bevarar materialet, medan ett mer tekniskt och positivistiskt orienterat synsätt kan föredra checksummor eftersom dessa visar att bitströmmens ursprungliga sekvensordning är oförändrad. Då de studerade institutionernas arkiv består av allmänna handlingar, vilket gör att målgruppen inte kan begränsas, medför det att ett helhetsperspektiv som inte prioriterar någon aspekt framför en annan ter sig mest fruktbar. OAIS-modellens metadatakartläggning sorterar metadata som behövs för bevarandet av digitala objekt på olika nivåer och ger en helhetsbild över vilken information som är nödvändig för att kunna uppnå långtidsbevarandet. Även om den inte är så detaljerad att den ger en checklista över vilka metadataelement som krävs för att ett digitalt objekt ska vara autentiskt kan den ge en fingervisning över hur väl sammansatt metadatan är eftersom alla metadataundergrupper bör inkluderas för att långtidsbevarandet ska kunna anses följa *best practice*. Att samtliga arkivinstitutioner uppfyller detta bådar gott för de framtida användarna. Som visades i analysen täcker de studerade institutionerna även in de autenticitetsaspekter som identifierats av diplomatiken som integritet och proveniens, samtidigt som deras egna rutiner och arbete dokumenteras vilket bidrar till transparensen i hur integriteten och proveniensen bevarats. Ur ett diplomatiskt perspektiv ter det sig därför som om att framtida användare, från det perspektiv som är tillgängligt nu, kommer att ha fullgoda möjligheter att bedöma autenticiteten hos de digitala objekten och att denna kommer att vara bibehållen. Detta konstateras dock med vissa avgörande reservationer: eftersom inga arkivformatsmigreringar genomförts är det omöjligt att sia om hur stor brytpunkt i integriteten dessa kommer att innebära. Dessutom gäller dessa förutsättningar för autenticitet *bara från det att arkivhandlingarna levererats till de studerade arkivinstitutionerna.*

Hur autenticiteten hos de digitala objekten behandlats under deras aktiva fas ute på myndigheter och förvaltningar kommer inte att framgå gällande merparten av arkivbestånden. Ser man originalet som den representation som ursprungligen uppstod i ärendehanteringssystemet då den allmänna handlingen först skapades är det

[77] Ett exempel för att förtydliga: En turist ska identifiera sig med sitt pass. Passkontrollanten, som har kunskap och erfarenhet av utseendet hos många länders pass, har bättre förutsättningar för att kontrollera handlingens autenticitet än nattklubbsvärden som aldrig tidigare sett ett pass från det landet. Passkontrollanten och nattklubbsvärden kan således ha olika uppfattningar om huruvida passet är autentiskt, men antingen är det utfärdat av turistens hemland eller inte och således äkta eller inte i absoluta termer. Detta är dock meningslöst för turisten om den nekas inträde till landet eller nattklubben på grund av att autenticiteten hos passet inte anses bevisad.

inte säkert att denna kommer att leveras till arkivinstitutionen eftersom okontrollerade och odokumenterade migreringar kan ha skett på myndigheterna vilket förvanskat denna. Om digitala förändringar inte dokumenteras är det mycket svårare att upptäcka och spåra dem än att spåra vissa förändringar i ett analogt arkivmaterial, även om inledningsexemplen visade att det kunde vara svårt nog. Precis som Lynch (2000, s. 36), Hirtle (2000, s. 10), Duranti (1995, s. 7; 2010, s. 80) och McNeil (1998, s. 1) tidigare konstaterat så innebär inte att ett *record* (eller digitalt objekt eller allmän handling heller för den delen) är autentiskt samma sak som att informationen denna innehåller är korrekt. Detta blir tydligt om man gör ett tankeexperiment där en förvaltning levererar digitala objekt som, antingen medvetet eller omedvetet, inte är vad de utges för att vara: exempelvis kan objekten ha märkts med fel metadata och personalhandlingar är märkta som ekonomihandlingar. Detta innebär att autenticitetsmarkören referensinformation är felaktig från förvaltningens sida, och att den autenticitet som förvaltningen förlänar innehållet ur den synpunkten inte kan anses som helt uppfylld. Hur uppfylld den är beror på hur man vill gradera de olika autenticitetsmarkörerna, något som inte gjorts i denna undersökning. Autenticitet är som bekant ett graderat värde vilket gör att den i det här fallet inte behöver vara totalt förlorad även om man som arkivarie instinktivt känner att ett objekts identitet är bland dess mest primära egenskaper. För att återgå till tankeexperimentet: förvaltningen har märkt de digitala objekten fel vid leverans, men det är inte arkivinstitutionen medveten om och felet upptäcks inte utan arkivet bevarar sedan leveransen på ett sätt som gör att dess fortsatta autenticitet är helt intakt. Är då de felmärkta handlingarna autentiska? Ja, från ett arkivvetenskapligt bevarandeperspektiv är de det eftersom arkivet bevarat dem så som de levererats till dem. Skulle en framtida användare hålla med? Tveksamt. Det är inte rimligt att arkivinstitutionerna ska garantera arbete som de inte själva utfört, oavsett om detta gäller bevarande eller ärendehantering, men de bör bevara förvaltningarnas dokumentation om detta arbete. Om inte förvaltningarna kan redovisa och garantera autenticiteten hos den information som skapats hos dem och som formaliserats som allmänna handlingar eller digitala objekt kommer framtida användare aldrig att kunna lita på det faktamässiga innehållet i arkivhandlingarna, bara att arkivhandlingarna i sig är autentiska. Att användaren kan lita på att faktan och informationsinnehållet i ett digitalt objekt är autentiskt är således en fråga om tilliten mellan användare och förvaltning. Tilliten mellan användaren och arkivet däremot påverkar huruvida användaren litar på att ingenting ändrats i det faktamässiga innehållet eller det informationsinnehåll som de digitala objekten är menade att återge.

I och med införandet av verksamhetsbaserad arkivredovisning kan kontexten, och därigenom proveniensen, för de allmänna handlingarna beskrivas mer detaljerat vilket påverkar autenticiteten såtillvida att de allmänna handlingarnas spårbarhet och arbetsprocessernas transparens ökar för de framtida användarna.

En av följderna för framtidens användare har att göra med originalets förlorade betydelse. I det analoga och fysiska originalet finns en existentiell dimension som inte existerar på samma sätt i en digital representation, även om denna också är autentisk. De flesta museibesökare har någon gång känt den historiska närhet och ödmjukhet som utställningsobjekt i form av unika historiska dokument kan ge. Den amerikanska självständighetsförklaringen, Magna Charta eller korrespondens mellan drottning Kristina och Descartes, återgivna på en datorskärm som en kopia och representation

av digitala bitströmmar hade de förlorat den omedelbara känslan av historicitet, artefaktens symbolik och mänsklig närvaro. Kanske kan detta påverka det gemensamma identitetsbyggande för vilket kulturarvet är så viktigt. Hur man kan återskapa denna existentiella dimension i artefakterna är möjligen snarast en fråga för museologin, men det hade varit högst intressant att ta del av resultaten även för de som sysslar med digitalt bevarande och arkivvetenskap. Samtidigt kan man spekulera i att en tillämpning av museimetoden för digitalt bevarande för denna typ av artefakters digitala bevarande och tillgängliggörande vore fruktbar eftersom det då kan skapas en historisk kontext runt själva representationen vilket fördjupar upplevelsen hos betraktaren. Som tidigare nämnts används just museimetoden av konstmuseer (Galloway 2004, s. 557).

Vilka behov och möjligheter för att använda digitala arkivhandlingar kommer då framtidens användare att ha? Det är, liksom informanterna menar, omöjligt att helt sia om även om det finns vissa trender som kan vara intressanta att nämna. En av trender som seglat upp på den informationsvetenskapliga horisonten de senaste åren representeras av begrepp som *big data* och *data mining*. Dessa koncept hamnade först i rampljuset då Google publicerade en artikel i Nature där de visade att genom att jämföra stora mängder data från sina sökförfrågningar med information från amerikanska smittskyddsinstitutiet om tidigare säsongsinfluensaspridningar kunde de förutse och påvisa spridning av influensa i realtid baserat på vilka ord som googlades (Mayer-Schönberger & Cukier 2013, s. 2). Man studerade alltså bara korrelationen mellan vilka ord som googlats och influensaförekomst, och försökte inte gissa eller pröva, vilket tidigare projekt gjort, vilka ord eller fraser som var kopplade till influensautbrott. *Big data* kan dock användas även inom andra fält; litteraturvetaren Jockers använde big data-konceptet till att göra en textanalys på 3 592 böcker publicerade mellan 1780–1900 för att undersöka vilka författare som varit de största språkliga och tematiska inspirationskällorna[78] under denna period (Lohr 2013). Jockers et al. (Jockers, Sag & Schultz 2012, s. 29) menar att digitala arkiv gör studier som grundar sig på *big data* möjliga även inom litteraturvetenskap och humaniora, även om copyrightproblematik kan sätta käppar i hjulet då upphovsrättsskyddat material måste digitaliseras för ändamålet. Det ter sig inte långsökt att även det arkivmaterial som de svenska offentliga arkiven innehåller skulle kunna användas för *data mining*. Kanske är framtidens historiker och samhällsvetare också programmerare? Frågan är dock hur man ska se till autenticiteten hos det färdiga resultatet av big data-analysen, hur ska den kunna bedömas när det, som denna uppsats visar, krävs så mycket för att bedöma autenticiteten hos ett enda dokument? Man kan spekulera i huruvida vissa av autenticitetsmarkörerna, särskilt de tekniska och automatiska, skulle kunna kvantifieras och presenteras i en sammanställning angående hela den stora datamängden. CASPAR-projektet har tagit fram en mjukvaruprototyp för att sammanställa och utvärdera autenticiteten hos enskilda arkivpaket genom att processera dess PDI (Giaretta 2011, ss. 221–227). Skulle andra hjälpmedel kunna utvecklas för att bygga vidare på detta och sammanställa autenticiteten hos större mängder arkivpaket? Då autenticitet hos det enskilda digitala

[78] Dessa visade sig vara Jane Austen och Sir Walter Scott. Resultaten publiceras i Jockers bok Macroanalysis: Digital Methods and Literary History.

objektet ses som något som kan graderas bör detta kunna överföras till större informationsvolymer, men hur detta kan ta sig uttryck eller vilka följder det får går inte att sia om inom ramen för denna uppsats. Dock är detta ett mycket intressant område för vidare forskning. Vidare forskning skulle också kunna bedrivas gällande huruvida autenticitetsmarkörerna som identifierats i denna uppsats kan användas för att kvantifiera autenticitet, både på enskilda digitala objekt och större aggregationer av dessa såsom arkivpaket, -bestånd eller större samlingar än detta, exempelvis på big datanivå. Hur autenticitetsmarkörerna skulle värderas gentemot varandra och hur utförligt de måste vara beskrivna för att kunna uppnå en godkänd miniminivå vore också intressant att utreda och vore en förutsättning för det förra forskningsinitiativet. Tillsammans skulle dessa forskningsinitivativ kunna ta fram, eller anpassa, programvara liknande CASPAR-prototypen för svenska juridiska förhållanden och där användarens syften och behov av materialets autenticitet tas med i beräkningen.

Annan vidare forskning inom detta område skulle kunna vara att undersöka hur myndigheterna och förvaltningarna som levererar till offentliga arkiv dokumenterar och bevarar autenticiteten hos de digitala objekten medan de är i deras vård. Förutom detta vore det intressant att studera hur de myndigheter som byggt upp mellanarkiv som inte kommer att levereras under lång tid hanterar autenticiteten.

Som tidigare diskuterats finns vissa tydliga skillnader mellan denna undersökning och tidigare forskning inom forskningsfältet. Denna undersökning kombinerar systemteori, OAIS och diplomatik på ett unikt sätt för att undersöka huruvida man kan se svenska offentliga arkivinstitutioner som system som tillhandahåller autentiska digitala allmänna handlingar. Undersökningen visar att det är möjligt att konstruera ett sådant system, och att samtliga undersökta arkivinstitutioner kan beskrivas enligt detta system, vilket i förlängningen visar att de har de rätta förutsättningarna för att bevara autentiska digitala allmänna handlingar. Genom att visa detta finns en möjlighet att undersökningen oavsiktligt bidrar till att öka samhällets tillit till arkivinstitutionerna, vilket i sin tur gör att ytterligare autenticitet förlänas arkivmaterialet genom tredjepartsintygandet. Även om man inte håller med om att arkivinstitutionerna bevarar autenticiteten i tillräckligt hög grad, eller accepterar de resultat som presenteras här, utgör undersökningen en inventering över vilka autenticitetsbevarande metoder som används inom svenska offentliga arkiv, vilket är ett resultat som aldrig tidigare sammanställts eller undersökts. Detta är av intresse inte bara för det arkivvetenskapliga forskningsfältet, utan kan också tjäna som en ögonblicksbild för framtida arkivanvändare som är intresserade av vad som ansågs vara *best practice*, och vilka metoder i det digitala bevarandet av autenticitet som användes, under det tidiga 2000-talet.

10.2 Metoddiskussion

Det kan tyckas okonventionellt att fenomenografi valdes som analysmetod eftersom den är en anpassning, som gjorts inom pedagogisk forskning, av fenomenologin (Kvale 1997, s. 56). Metoden har förvisso fått större spridning och har även använts inom andra forskningsfält, exempelvis biblioteks- och informationsvetenskap. Den stora skillnaden mellan fenomenografi och fenomenologi är att den förra inte är beroende av de djuplodade filosofiska resonemang som är kännetecknande för den senare. Båda analysmetoderna är framtagna för att användas inom kvalitativ

forskning för att tolka utsagor som ett led i att undersöka hur individer uppfattar omvärlden och bygger kunskap om och tolkar denna (Patton 2002, s. 104). För att besvara forskningsfrågan i denna undersökning räcker inte en renodlat kvalitativ utgångspunkt eftersom det även behövs ett positivistiskt perspektiv för att besvara och kategorisera de olika autenticitetsmarkörerna. På grund av denna tudelning i perspektiv är det mer gynnsamt för undersökningen att systemteorin är det överordnade teoretiska perspektivet snarare än fenomenografin. Därför sågs det som gynnsamt att endast använda den fenomenografiska analysproceduren som ett verktyg för att tematisera och tolka resultaten, snarare än att låta den och dess systermetod genomsyra hela det metodologiska och teoretiska perspektivet.

Kombinationen av fenomenografi som analysmetod med intervjuer och litteraturgranskning som ett sätt att samla in det empiriska materialet föll väl ut, och är ett etablerat sätt att använda analysmetoden (Dahlgren & Johansson 2009, s. 122). Att kombinera fenomenografi och systemteori var särskilt gynnsamt eftersom fenomenografisk analysmetod syftar till att tematisera och dela upp det empiriska materialet i exklusiva kategorier. Då systemteori bygger på att delar eller sammankopplingar inom ett system växelverkar med varandra på ett ickekausalt sätt kompletterade detta teoretiska perspektiv och fenomenografin som metod varandra väldigt väl. Då kategorier, sammankopplingar och element alla utgör avgränsade delar av en helhet kunde fenomenografin användas för att identifiera autenticitetsmarkörerna som sedan kunde tolkas genom systemteorin och placeras in i systemet.

De teoretiska och metodologiska perspektiven som använts i undersökningen var särskilt gynnsamma för att besvara forskningsfrågan av flera anledningar. Den första delen av forskningsfrågan, "vilka autenticitetsmarkörer används av offentliga arkivinstitutioner", är positivistisk och deskriptiv till sin natur och besvarades därför väl med hjälp av fenomenografins kategoriseringsmetodik. Tolkningen av denna kategoriering genomfördes därefter med det teoretiska perspektivet på flera plan. Dels genom att de olika autenticitetsmarkörerna kunde tolkas som rörande proveniens/identitet, integritet eller en kombination av båda faktorerna genom autenticitetssynen från OAIS och diplomatik, men även genom att autenticitetsmarkörerna kunde tolkas och placeras in i systemet med hjälp av systemteorin. I denna tolkning framkom det exempelvis att det var rimligare att knyta bevarandeåtgärder och tidsstämplar till respektive autenticitetsmarkör istället för att se bevarandehistorik som en unik kategori. "Varför har dessa metoder/metadata valts ut och vilka konsekvenser får detta urval för de framtida användarna" från forskningsfrågan besvaras också på ett fördelaktigt vis genom fenomenografi och systemteori. Varför vissa autenticitetsmarkörer valts ut kunde härledas ur de förutsättningar för autenticitetsbevarande som uppstått då migrering valts som bevarandestrategi. Denna tolkning underlättades och möjliggjordes av fenomenografin genom dess kategorisering av autenticitetsmarkörerna som kopplade till särskilda bevarandestrategier. Förutom detta kunde valet av markörer identifieras som beroende av autenticitetssynen som i sin tur är beroende av migreringsstrategin

eftersom denna påverkar vilka aspekter av autenticiteten hos ett digitalt objekt som kan vara föränderliga. Uppfattningen att synen på ett fenomen påverkar individens, eller i det här fallet institutionens,[79] agerande och omvärldsuppfattning härstammar från fenomenografin och fenomenologin. Samtidigt är bevarandestrategins påverkan på autenticitetsmarkörerna och autenticiteten ett utslag av ett ickekausalt del – helhetstänkande hämtat från systemteorin. Således kompletterar systemteorin och fenomenografin varandra genom att fenomenografin identifierar delar som kan inkorporeras i systemets helhet av systemteorin.

Då valet av inklusionskriterier begränsade de studerade objekten till mellan- och slutarkiv valdes möjligheten bort att studera digitala objekt från deras skapelse till slutförvaring. Detta var ett medvetet val eftersom redan under arbetet med projektplanen det blev tydligt att förutsättningarna för att studera autenticitet på förvaltningar respektive på arkiven var radikalt annorlunda. Riksarkivets enkätundersökning till samtliga statliga myndigheter 2010 visade att det digitala bevarandet på en betydande del av myndigheterna var beroende av ad hoc-lösningar och inte alltid gjordes under de bästa förhållandena (Jarborn & Gäfvert 2010). InterPARES 1 (InterPARES 2002b) identifierade en tydlig brytpunkt mellan den aktiva och de semi- och inaktiva faserna i records lifecyclemodellen och studerade därför endast mellan- och slutarkiven, med tämligen goda resultat. Hänström (2007, s. 83) menade att det även för svenska förhållanden fanns tendenser att autenticitet var något som först började tas i beaktande vid leverans till arkivet. Detta sammantaget med att svårigheterna i att få ett representativt urval av myndigheter[80] och att undersökningen hade blivit för omfattande för en mastersuppsats gjorde att beslutet fattades att inte inkludera den aktiva fasen. Dock finns det styrkor i att följa ett dokument från skapelsen till arkivet, och detta hade kunnat ge bättre svar på vilka möjligheter användaren får för att spåra autenticiteten. Gör man detta integrerar man ett records managementperspektiv med det rent arkivvetenskapliga, för som Hirtle (2000, s. 10) menar så är det inte arkivens sak att garantera autenticiteten för mer än det de får in. Dock visade sig i undersökningen att även i de fall dokumentation om bevarandet och autenticiteten hos de digitala dokumenten upprättats på institutionerna bevarades den inte alltid. Med det resultatet i åtanke är det därför osäkert om ett studium av även den aktiva fasen tillfört någonting eftersom resultatet från den aktiva fasen lika gärna kunde negeras och förlora betydelse om den inte kunde fortlöpa in i nästkommande fas.

Det hade kunnat vara önskvärt att undersökningen omfattat fler arkivinstitutioner, tyvärr uppstod svårigheter att få tag på informanter. Dessutom gjordes bedömningen att undersökningen skulle lida av att vara asymmetrisk med fler studerade arkivinstitutioner av en kategori än en annan. Att de tre studerade institutionerna använde sig av så likartad metodik kan dock tyda på att skillnaderna sinsemellan

[79] Institutionen är en grupp av individer och enligt Patton (2002, s. 104) kan fenomenologi användas på både individ och gruppnivå.
[80] Riksarkivets enkätundersökning bestod av 344 enkätsvar från statliga myndigheter av det ursprungliga utskicket på 387 enkäter vilket visar storleksordningen på populationen som urvalet skulle gjorts ur.

olika offentliga e-arkiv generellt sett inte är så stora och att undersökningen började uppnå eller uppnått datamättnad.

Gällande insamlandet av det empiriska materialet är det lämpligt att diskutera huruvida andra metoder än intervjuer hade varit att föredra. Alvesson och Sköldberg (1994, s. 126) menar att enligt hermeneutisk källkritik är protokoll mer värda som källor än intervjuuttalanden delvis för att de ligger närmre själva händelsen de beskriver men också för att de kan innehålla mer tendens. Ursprungligen var tanken att även metadataschema från *alla* de studerade institutionerna skulle lägga grunden för det empiriska materialet och endast kompletteras med intervjuer för att nå de kvalitativa insikter som metadataschemana allena inte kunde ge. Det visade sig dock vara svårt att få tag på likvärdigt material från samtliga institutioner, vilket gjorde att ett beslut fattades att den huvudsakliga tyngdpunkten för resultat och analys skulle ligga på intervjumaterialet för att undersökningsdjupet skulle bli jämnare på samtliga institutioner. Detta var ett fördelaktigt beslut också på så vis att de rapporter och metadataschema som kunde studeras var mindre rika på information än vad som först antagits, vilket gjorde att de i realiteten bäst passade som underlag för fördjupade intervjufrågor snarare än som empiriskt material i sig själva. Det kunde ha varit att föredra en kombination av granskning av metadatascheman och intervjuer på samtliga arkivinstitutioner för att dels få lättare att jämföra institutionerna sinsemellan men även för att få tydligare evidens för vilka val de olika institutionerna gjort. Det visade sig dock vara svårt att få tillgång till metadataschemana eftersom de var under uppbyggnad eller omarbetning i merparten av fallen. Dock hade informanterna stora kunskaper om metadataschemana och deras uppbyggnad, och i de fall de inte kunde besvara frågor om specifika metadataelement under intervjutillfället återkom de med svar efter att de granskat dokumentationen eller förhört sig med kollegor inom institutionen. Således är det ett rimligt antagande att kvaliteten hos det empiriska materialet blev tillräckligt högt för att ligga till grund för valida slutsatser.

11. Slutsats

Föreliggande undersökning tar sitt avstamp i den övergripande forskningsfrågan: ***Vilka autenticitetsmarkörer för att bevara autenticitet hos digitala objekt används av offentliga arkivinstitutioner, varför har dessa metoder/metadata valts ut och vilka konsekvenser får detta urval för de framtida användarna?***

De tre undersökta arkivinstitutionerna använder en rad autenticitetsmarkörer i beskrivandet av sina digitala objekt. När det gäller objektets integritet används checksummor och formatvalidering medan proveniensen beskrivs av objektets identitet, ursprung, kontext och ansvarskedja. Detta kombineras med bevarandet av system- och arbetsdokumentation. Transparensen och dokumentationen gällande arbetsrutiner, processer och beslut ger goda förutsättningar för allmänhetens fortsatta tillit gentemot arkivinstitutionerna. Dels för att det ger en utförlig och heltäckande beskrivning av det digitala objektet, dels för att det visar att arkivinstitutionen, i alla fall delvis, arbetar efter dagens best practice-metoder vilket ökar tilltron till arkivhandlingens autenticitet enligt principen om tredjepartsintygande. Detta resultat har genom systemteori tolkats och åskådliggjorts i ett system över det offentliga svenska e-arkivet där elementet information förädlas genom autenticitetsmarkörernas sammankopplingar för att resultera i funktionen autentiska digitala allmänna handlingar.

Att de studerade arkivinstitutionernas metodik för att säkra autenticitet och de autenticitetsmarkörer som då skapas är så lika varandra kan tolkas som ett resultat av att de har likartade krav ställda på sig från omgivningen, men också att det de facto inte finns så många olika metoder att välja mellan. Internationellt beprövade tekniska metoder som checksummor kombineras med metadatabeskrivningar enligt traditionell arkivteori och diplomatik. Detta sammantaget medför att *best practice* för institutionerna blir desamma. Det kan också ses som att det var adekvat att med systemteoretiskt perspektiv se olika offentliga arkivinstitutioner som system som fungerar på samma sätt avseende autenticitet hos digitala objekt.

De migreringar av de digitala objekten som med säkerhet kommer att utföras ett antal gånger i det längre perspektivet kommer att bli en brytpunkt för autenticiteten. Det är svårt att göra några detaljerade förutsägelser om hur detta kommer att ske då man inte vet hur de framtida filformaten eller mjukvaran kommer att se ut, informanterna spekulerar även ogärna om formerna för detta. Klart är att man vid migreringar måste ta ställning till vad det egentligen är som ska bevaras för att säkerställa väsentlig information, vilket gör att institutionerna för tillfället är öppna för möjligheten att använda sig av metodiken kring signifikanta egenskaper för att säkra autenticiteten och bevarandet. Följden av en felaktig migrering kan bli ofrivillig gallring. Å andra

sidan ska inte faran överdrivas, migreringar kan göras om, men väntar man för länge riskeras obsolescens och sönderfall.

Som förutsättningarna ser ut nu, det vill säga att inga arkivformatmigreringar gjorts, kan alltså framtida användare komma att ha goda förutsättningar att bedöma autenticiteten hos digitala allmänna handlingar under förutsättning att den dokumentation som skapas kring de digitala objekten håller hög kvalitet. Att alla studerade arkivinstitutioner använder samtliga PDI-metadataundergrupper i OAIS-modellens metadataöversikt är dubbelt gynnsamt för användaren eftersom det både ger förutsättningar för att användaren får en heltäckande bild av autenticiteten och tyder på att institutionen arbetar enligt *best practice*. Dock är det omöjligt att uttala sig generellt om autenticiteten hos det enskilda digitala objektet eftersom det beror på kvaliteten på informationen i autenticitetsmarkörerna som beskriver den.

12. Referenslista

Adam, S. (2010). Preserving authenticity in the digital age. *Library Hi Tech*, vol. 28:4, ss. 595–604.

Alvesson, M. & Sköldberg, K. (1994). *Tolkning och reflektion: vetenskapsfilosofi och kvalitativ metod*. Lund: Studentlitteratur.

Askergren, K. (2009). Långsiktighet och ansvar vid arkivering av pensionsgrundande uppgifter. I Näringslivets arkivråd (red.) *E-arkivera rätt: sju perspektiv på hantering av digital information med hjälp av OAIS.* Stockholm: Näringslivets arkivråd.

Backman, J. (1998). *Rapporter och uppsatser*. Lund: Studentlitteratur.

Belton, T. (1996). By whose Warrant? Analyzing Documentary Form and Procedure. *Archivaria,* vol. 41, ss. 206–220.

Bohlin, A. (2010). *Offentlighetsprincipen*. 8. uppl. Stockholm: Norstedts Juridik.

Boudrez, F. (2007). Digital signatures and electronic records. *Archival Science,* vol. 7, ss. 179–193.

Bradley, R. (2005). Digital Authenticity and Integrity: Digital Cultural Heritage Documents as Research Resources. *portal: Libraries and the Academy,* vol. 5:2, ss. 165–175.

Brown, A. (2006). *Digital Preservation Paper 1: Automatic Format Identification Using PRONOM and DROID.* London: The National Archives. Tillgänglig: http://www.nationalarchives.gov.uk/aboutapps/fileformat/pdf/automatic_format_identification.pdf [2013-03-28]

Buckland, M. (1997). What is a 'Document'? *Journal of the American Society for Information Science*, vol. 48:9, ss. 804–809. Tillgänglig: http://people.ischool.berkeley.edu/~buckland/whatdoc.html [2013-03-18]

Buckland, M. (1998). What is a Digital Document? *Document Numérique* (Paris), vol. 2:2, ss. 221–230. Tillgänglig: http://people.ischool.berkeley.edu/~buckland/digdoc.html [2013-03-18]

The Consultative Committee for Space Data Systems, CCSDS. (2012). *Reference model for an Open Archival Information System (OAIS). Recommended Practice CCSDS 650.0-M-2. Magenta book.* Washington DC: The Consultative Committee for

Space Data Systems Secretariat. Tillgänglig: http://public.ccsds.org/publications/archive/650x0m2.pdf [2013-02-26]

Conway, P. (2010). Preservation in the age of Google: digitization, digital preservation, and dilemmas. *Library Quarterly*, vol. 80:1, ss. 61–79.

Cox, R. & Archives Students, The. (2007). Machines in the archives: Technology and the coming transformation of archival reference. *First Monday*, vol. 12:11. Tillgänglig: http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2029/1894 [2013-03-20]

Cullen, C. (2000). Authentication of Digital Objects: Lessons from a Historian's Research. I Council on Library and Information Resources, CLIR (red.) *Authenticity in a Digital Environment.* (Rapport). Washington DC: Council on Library and Information Resources. Tillgänglig: http://www.clir.org/pubs/reports/pub92/reports/pub92/pub92.pdf [2013-03-20]

Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval, CASPAR. (2006). *The CASPAR project.* Tillgänglig: http://www.casparpreserves.eu/caspar-project.html [2013-04-28]

Dahlgren, L. O. & Johansson K. (2009). Fenomenografi. I Fejes, A. & Thornberg, R. (red.) *Handbok i kvalitativ analys.* 1:a uppl. Stockholm: Liber.

Delmas, B. (1996). Manifesto for a Contemporary Diplomatics: From Institutional Documents to Organic Information. *American Archivist*, vol. 4, ss. 439–445.

Digital Preservation Testbed Project (2003). *Emulation: Context and current status.* The Hague: ICTU. Tillgänglig: http://www.google.se/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&sqi=2&ved=0CDYQFjAB&url=http%3A%2F%2Fciteseerx.ist.psu.edu%2Fviewdoc%2Fdownload%3Bjsessionid%3DD13CBBA3AA40BA98F5511427D073DE47%3Fdoi%3D10.1.1.132.5566%26rep%3Drep1%26type%3Dpdf&ei=A_GlUabTEYj_4QSHpICoDw&usg=AFQjCNEFxIMS8zCGIlM1llkO9ZNWpAmaJw&sig2=Fd4Cpp7bqKY5FNLftlyTLw&bvm=bv.47008514,d.bGE [2013-05-30]

Dollar, C. (2000). *Authentic electronic records: strategies for long-term access*. Chicago, Ill.: Cohasset Associates, Inc.

Duff, W. & Harris, V. (2002). Stories and Names: Archival Description as Narrating Records and Constructing Meanings. *Archival Science*, vol. 2, ss. 263–285.

Duranti, L. (1989). *Diplomatics: new uses for an old science*. Lanham, Md.: Scarecrow Press.

Duranti, L. (2008). Introduction. I Duranti, L. & Preston, R. (red.) *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records.* Padova: Associazione Nazionale

Archivistica Italiana. Tillgänglig: http://www.interpares.org/ip2/display_file.cfm?doc=ip2_book_complete.pdf [2013-03-31]

Duranti, L. (2010). Concepts and principles for the management of electronic records, or management theory is archival diplomatics. *Records Management Journal*, vol. 20:1, ss. 78–95.

Eastwood, T. (1994). What is archival theory and why is it important? *Archivaria*, vol. 37, ss. 122–130.

e-Arkiv och e-Diarium, eARD (2013). *Begreppslista version 1.1*. (2013-01-14). Tillgänglig: www.riksarkivet.se/default.aspx?id=27903&ptid=0 [2013-05-30]

Eriksson-Zetterquist, U. & Ahrne, G. (2011) Intervjuer. I Ahrne, G. & Svensson, P. *Handbok i kvalitativa metoder.* 1. uppl. Malmö: Liber.

Faniel, I. & Yakel, E. (2011). Significant Properties as Contextual Metadata. *Journal of Library Metadata*, vol. 11, ss. 155–165.

Fredriksson, B. (2003). Vad skall vi bevara? *Arkiv, samhälle och forskning*, vol. 2, ss. 21–58.

Frendo, R. (2007). Disembodied information: Metadata, file plans, and the intellectual organization of records, *Records Management Journal*, volym 17:3, ss. 157–168.

Giaretta, D. (2011). *Advanced Digital Preservation*. [Elektronisk]. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg. Tillgänglig: Springer Online. [2013-03-19]

Galloway, P. (2004). Preservation of Digital Objects. *Annual Review of Information Science and Technology*, vol. 38:1, ss. 549–590.

Gilliland-Swetland, A. (2005). Electronic Records Management. *Annual Review of Information Science and Technology*, vol. 39, s 219-253.

Gladney, H. (2007). *Preserving digital information.* [Elektronisk]. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg. Tillgänglig: Springer Online. [2013-03-19]

Gränström, C., Hornfeldt, T., Peterson, G., Rinaldi Mariana, M. P., Schäfer, U. & Zwicker, J. (2002). *Authenticity of electronic records: A Report prepared for UNESCO. ICA Study 13-1*. [Elektronisk] Paris: International Council of Archives. Tillgänglig: www.**ica**.org/download.php?id=1624 [2013-03-20]

Gränström, C., Lundquist, L. & Fredriksson, K. (2000). *Arkivlagen: bakgrund och kommentarer*. 2. uppl. Stockholm: Norstedts juridik.

Harvard Business School (2008). Niall Ferguson lectures on globalization. [Föreläsning] http://www.youtube.com/watch?v=VqXQP3eB9W0 [2012-12-17]

Hermerén, G. (2011). *God forskningssed*. Stockholm: Vetenskapsrådet. Tillgänglig: http://www.vr.se/download/18.3a36c20d133af0c12958000491/1321864357049/God+forskningssed+2011.1.pdf [2013-04-14]

Heslop, H., Davis, S. & Wilson, A. (2002). *An Approach to the Preservation of Digital Records.* [Elektronisk] Canberra: National Archives of Australia. Tillgänglig: http://www.naa.gov.au/Images/An-approach-Green-Paper_tcm16-47161.pdf [2013-03-18]

Hessler, G. (2003). *Identitet och förändring: en studie av ett universitetsbibliotek och dess självproduktion.* Diss. Bibliotekshögskolan/Högskolan i Borås och Göteborgs universitet. Tillgänglig: http://bada.hb.se/bitstream/2320/2401/1/GunnelHessleravhandling.pdf [2013-04-29]

Hirtle, P. (2000). Archival authenticity in a digital age. I Council on Library and Information Resources, CLIR (red.) *Authenticity in a Digital Environment.* (Rapport). Washington DC: Council on Library and Information Resources. Tillgänglig: http://www.clir.org/pubs/reports/pub92/reports/pub92/pub92.pdf [2013-03-20]

van der Hoeven, J. & van Wijngaarden, H. (2005). *Modular emulation as a long-term strategy for digital objects.* The Hague: Koninklijke Bibliotheek. Tillgänglig: http://scholar.google.se/scholar_url?hl=en&q=http://citeseerx.ist.psu.edu/viewdoc/download%3Fdoi%3D10.1.1.106.2100%26rep%3Drep1%26type%3Dpdf&sa=X&scisig=AAGBfm1pX68ajTaD8ekF0zD797XjVrISIw&oi=scholarr&ei=A_GlUabTEYj_4QSHpICoDw&sqi=2&ved=0CCsQgAMoAjAA [2013-05-28]

Hänström, K. (2007). Autenticitet i en digital värld: långsiktsbevarande av allmänna handlingar. *Human IT,* vol. 9:1, ss. 67–109.

Ilshammar, L. (2008). Arkiven och den digitala paradoxen. I Bergman, Y., Eriksson, B. & Fransson, B. (red.) *Titta vad vi har! Nedslag i de enskilda arkiven*. Örebro: Folkrörelsernas arkivförbund. Tillgänglig: http://ilshammar.se/wp-content/uploads/2008/10/nordisk-klarsprakskonferens.pdf

International Council on Archive, ICA, Committee on Electronic Records. (1997). *Guide for Managing Electronic Records from an Archival Perspective.* (ICA study number 8). Paris: International Council of Archives. Tillgänglig: http://www.ica.org/10824/studies-and-case-studies/ica-study-n8-guide-for-managing-electronic-records-from-an-archival-perspective.html [2013-03-18]

International Research on Permanent Authentic Records in Electronic Systems, InterPARES 1. (2000). Appendix 1: Template for Analysis. I International Research on Permanent Authentic Records in Electronic Systems, InterPARES 1. (red.) *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project.* Vancouver: InterPARES.

Tillgänglig: http://www.interpares.org/book/interpares_book_j_app01.pdf [2013-03-31]

International Research on Permanent Authentic Records in Electronic Systems, InterPARES 1. (2002a). Appendix 2: Requirements for Assessing and Maintaining the Authenticity of Electronic Records. I International Research on Permanent Authentic Records in Electronic Systems, InterPARES 1. (red.) *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project.* Vancouver: InterPARES. Tillgänglig: http://www.interpares.org/book/interpares_book_k_app02.pdf [2013-03-31]

International Research on Permanent Authentic Records in Electronic Systems, InterPARES 1. (2002b). Introduction. I International Research on Permanent Authentic Records in Electronic Systems, InterPARES 1. (red.) *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project.* Vancouver: InterPARES. Tillgänlig: http://www.interpares.org/book/interpares_book_c_intro.pdf [2013-03-31]

International Research on Permanent Authentic Records in Electronic Systems, InterPARES 1. (2002c).
Authenticity Task Force Report. I International Research on Permanent Authentic Records in Electronic Systems, InterPARES 1. (Red.) *The Long-term Preservation of Authentic Electronic Records: Findings of the InterPARES Project.* Vancouver: InterPARES. Tillgänglig: http://interpares.org/display_file.cfm?doc=ip1_atf_report.pdf [2013-03-31]

Investigating SIgnificant Properties of Electronic Content, InSPECT. (2012). *Significant Properties and Digital Preservation.* Tillgängligt: http://www.significantproperties.org.uk/ [2013-04-13]

Jarborn, E. & Gäfvert, T. (2010). *Rapport rörande enkätundersökningar: myndigheters hantering av elektroniska handlingar.* (Rapport). Stockholm: Riksarkivet. Tillgänglig: http://www.riksarkivet.se/Sve/Dokumentarkiv/Filer/enkatundersokning_e-handlingar_rapport.pdf [2013-04-27]

Jockers, M., Sag, M. & Schultz, J. (2012). Digital archives: Don't let copyright block data mining. *Nature,* vol. 490, ss. 29–30.

Johnson, A. (2008). Users, Use and Context: Supporting Interaction Between Users and Digital Archives. I Craven, L. (red.) *What are Archives? Cultural and Theoretical Perspectives: A Reader.* Aldershot: Ashgate.

Katz, G. (2008). Rewriting History in Great Britain. *Smithsonian Magazin*e. [Elektronisk] 18 november. Tillgänglig: http://www.smithsonianmag.com/history-archaeology/Rewriting-History-in-Great-Britain.html [2012-12-17]

Kvale, S. (1997). Den kvalitativa forskningsintervjun. Lund: Studentlitteratur.

Langefors, B. (1993). *Essays on Infology: Summing up and planning for the future.* Göteborg: Department of Information Systems, University of Göteborg.

Levy, D. (2000). Where's Waldo? Reflections on Copies and Authenticity in a Digital Environment. I Council on Library and Information Resources, CLIR (red.) *Authenticity in a Digital Environment.* (Rapport). Washington DC: Council on Library and Information Resources. Tillgänglig: http://www.clir.org/pubs/reports/pub92/reports/pub92/pub92.pdf [2013-03-20]

Library of Congress (2008). JHOVE and the Development of JHOVE2. *Digital Preservation.* Tillgänglig: http://www.digitalpreservation.gov/news/2008/20080902news_article_JHOVE2.html [2013-05-25]

Lohr, S. (2013). Dickens, Austen and Twain, Through a Digital Lens. *The New York Times,* 26 januari. Tillgänglig: http://www.nytimes.com/2013/01/27/technology/literary-history-seen-through-big-datas-lens.html?pagewanted=all&_r=1& [2013-04-26]

Lundahl, C. & Öquist, O. (2002). *Idén om en helhet: utvärdering på systemteoretisk grund*. Lund: Studentlitteratur.

Lynch, C. (2000). Authenticity and Integrity in the Digital Environment: An Exploratory Analysis of the Central Role of Trust. I Council on Library and Information Resources, CLIR (red.) *Authenticity in a Digital Environment.* (Rapport). Washington DC: Council on Library and Information Resources. Tillgänglig: http://www.clir.org/pubs/reports/pub92/reports/pub92/pub92.pdf [2013-03-20]

Långsiktigt digitalt bevarande-centrum, LDB-centrum (2013). *Om LDB-centrum.* Tillgängligt: http://www.ltu.se/centres/Centrum-for-langsiktigt-digitalt-bevarande-LDB/Om-oss [2013-05-20]

Lövblad, H. (2004). Monk, Knight or Artist? The Archivist as a Straddler of a Paradigm. *Archivaria,* vol. 3, ss. 131–155.

Mayer-Schönberger, V. & Cukier, K. (2013). *Big data: a revolution that will transform how we live, work, and think.* Boston, Mass.: Eamon Dolan/Houghton Mifflin Harcourt.

McAfee, A. & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review,* oktober. Tillgänlig: http://hbr.org/2012/10/big-data-the-management-revolution/ar [2013-05-20]

McNeil, H. (1998). *Trusting Records: The Evolution of Legal, Historical, and Diplomatic Methods of assessing the Trustworthiness of Records, from Antiquity to the Digital Age.* Diss. The University of British Columbia. Tillgänglig: https://circle.ubc.ca/handle/2429/10157

Meadows, D. (2008). *Thinking in systems: a primer*. Vermont: Chelsea Green Pub.

Merriam, S. (1994). *Fallstudien som forskningsmetod*. Lund: Studentlitteratur.

Moss, M. (2008). Opening Pandora's Box: What is an Archive in the Digital Environment? I Craven, L. (red.) *What are Archives?: Cultural and Theoretical Perspectives: a Reader*. Aldershot: Ashgate.

The National Archives. (uå) *Investigation into Forged Documents discovered amongst Authentic....* Tillgänglig: http://discovery.nationalarchives.gov.uk/SearchUI/details?Uri=C16525 [2012-12-17]

The National Archives and Records Administration, NARA. (2011). *National Archives Discovers Date Change on Lincoln Record.* Pressrelease. Washington D.C.: NARA. Tillgänglig: http://www.archives.gov/press/press-releases/2011/nr11-57.html [2012-12-17]

Nilsson, J. (2008). *Preserving Useful Digital Obejcts for the Future.* Diss. Luleå University of Technology.

Nilsson, J. & Hägerfors, A. (2007). Metadata Driven Presentation of Digital Documents/Records. I Stillman, L. & Johanson, G. (red.) *Constructing and sharing memory: community informatics, identity and empowerment*. Newcastle, UK: Cambridge Scholars Publishing.

Patton, M. Q. (2002). *Qualitative research & evaluation methods*. 3. ed. London: SAGE.

Preservation and Long-term Access through Networked Services, PLANETS. (2009). *PLANETS: Tools and Services for Digital Preservation*. Tillgänglig: http://www.planets-project.eu/docs/comms/PLANETS_PRODUCT_SPECIFICATION.pdf [2013-03-25]

Quisbert, H. (2008). *On Long-term Digital Preservation Information Systems. A Framework an Characteristics for Development.* Diss. Luleå University of Technology.

Riksarkivet. (2011). *Förstudie om e-arkiv och e-diarium.* (Rapport). Stockholm: Riksarkivet Tillgänglig: http://www.riksarkivet.se/Sve/Dokumentarkiv/Filer/Forstudie_e-arkiv_e-diarium_Rapport.pdf [2013-03-05]

Roeder, J., Eppard, P., Underwood, W. & Lauriault, T. (2008). Part Three: Authenticity, Reliability and Accuracy of Digital Records in the Artistic, Scientific and Governmental Sectors: Domain 2 Task Force Rapport. I Duranti, L. & Preston, R. (red.) *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records.* Padova: Associazione Nazionale Archivistica Italiana. Tillgänglig: http://www.interpares.org/ip2/display_file.cfm?doc=ip2_book_complete.pdf [2013-03-31]

Rothenberger, J. (2000). Preserving Authentic Digital Information. I Council on Library and Information Resources, CLIR (red.) *Authenticity in a Digital Environment.* (Rapport). Washington DC: Council on Library and Information Resources. Tillgänglig: http://www.clir.org/pubs/reports/pub92/reports/pub92/pub92.pdf [2013-03-20]

Runardotter, M., Quisbert, H., Nilsson, J., Hägerfors, A., & Mirijamdotter, A. (2006). The Information Life Cycle: Issues in Long-term Digital Preservation. *Arkiv, samhälle och forskning,* vol. 1, ss. 17–29.

Sahlén, T. (2005). Kaos eller struktur: om modern dokumenthantering. I Sundquist, A. (red.) *Dokumenthantering i processorienterade organisationer.* Stockholm: Näringslivets Arkivråd (NLA).

Samson, F. (2009). De funktionella komponenterna i ett e-arkiv enligt OAIS. I Näringslivets arkivråd (red.) *E-arkivera rätt: sju perspektiv på hantering av digital information med hjälp av OAIS.* Stockholm: Näringslivets arkivråd.

Shepherd, E. & Yeo, G. (2003). *Managing Records: a handbook of principles and practice.* London: Facet Publishing.

Storch, S. (1998). Diplomatics: Modern Archival Method or Medieval Artifact. *American Archivist,* vol. 2, ss. 381–383.

Sustaining Heritage Access through Multivalent ArchiviNg, SHAMAN (uå.) *Scientific & Technical Goals.* Tillgänglig: http://shaman-ip.eu/node/8 [2013-04-15]

Svensson, P. & Ahrne, G. (2011). Att designa ett kvalitativt forskningsprojekt. I Ahrne, G. & Svensson, P. *Handbok i kvalitativa metoder.* 1. uppl. Malmö: Liber.

Utvik, M. (2006). En strategisk modell för hantering av autentiska elektroniska handlingar. *Arkiv, samhälle och forskning,* vol. 1, ss. 7–16.

Yeo, G. (2010). 'Nothing is the same as something else': significant properties and notions of identity and originality. *Archival Science,* vol. 10, ss. 85–116.

ZD Net Australia. (2002). *Teach Guide: Storage media lifespan*. Tillgänglig: http://www.zdnet.com.au/reviews/hardware/storage/0,39023427,20269043,00.htm [2013-03-14]

Öberg L.-M. & Borglund, E. (2006). What are the Characteristics of Records? *Arkiv, samhälle och forskning,* vol 1, ss. 30–48.

Bilaga 1: Intervjufrågor

1. Beskriv det digitala arkivet: Storlek (hur stora datamängder, hur många delarkiv) Hur länge har e-arkivet varit i drift? Hur många arbetar med e-arkivet?

2. Är e-arkivet uppbyggt efter OAIS-modellen?

3. Har migrering/transformering av stora datamängder varit nödvändigt?
Beskriv gärna hur detta gick till; hur såg beslutsprocessen ut, vilka rutiner fanns redan eller utarbetades under arbetets gång. Kunde migreringen/transformeringen genomföras enligt planerna? Hur undersöktes om representationen av handlingarnas innehåll bibehållits?

4. Har praktiska problem med autenticitet eller integritet uppstått?

5. Hur kontrollerar ni att e-arkivets handlingar behåller sin integritet under lagring?

6. Använder ni samma metadatascheman för hantering av material som är scannat respektive born-digital? Text respektive bilder? Sammansatta respektive enkla digitala objekt?

7. Vid framtagandet av bevarandestrategier, i vilken mån diskuteras framtida användningsområden och användare?

8. Hur fungerar kontinuiteten för exempelvis proveniens och annan metadata kopplad till autenticitet vid leverans till e-arkivet?

9. Hur ser Du på:
Originalitet i elektroniska/analoga dokument?
Autenticitet och integritet hos elektroniska dokument?
Bevarandet av layout hos ett dokument jämfört med att bevara datan det innehåller?

10. Ser Du några vidare utmaningar gällande autenticitet annat än det vi talat om?

Bilaga 2: Presentationsbrev till institutionerna

Hej!

Vi är två studenter som studerar arkivvetenskap på masternivå på Lunds universitet. Under våren arbetar vi med vår mastersuppsats som behandlar autenticitet inom digitala arkiv. Det är en kvalitativ studie som är tänkt att grundas i intervjuer av personal med kunskap om e-arkiv på offentliga arkivinstitutioner, samt analys av metadataschemana för dessa e-arkiv. Uppsatsens frågeställning inbegriper ett användarperspektiv där vi tänkt analysera vilka möjligheter medborgaren, forskaren eller tjänstemannen om hundra år har att bedöma autenticiteten hos de digitala dokument som skapas i vår tid.

Vår fråga till Er är om Ni, eller någon på er institution med insikt i frågan, skulle vilja ställa upp på en intervju i detta ämne? Vi är även intresserade av att få ta del av Era metadatascheman. Om Ni skulle vilja ställa upp på detta, vilket hade glatt oss storligen, så får Ni gärna e-posta eller ringa någon av oss. Intervjuerna är tänkta att ske i slutet av februari eller i första halvan av mars. Beräknad tidsåtgång är högst sextio minuter.

Vi vill gärna spela in intervjuerna på band, dels som en hjälp för oss att komma ihåg vad som sagts, men även som en säkerhet för vår informant så att hon/han blir korrekt återgiven. Vidare har vi planerat att strukturera intervjun efter ett antal uppställda frågor. Om Ni vill ha frågorna i förväg går det utmärkt. Om informanten önskar vara anonym i uppsatsen avidentifierar vi naturligtvis denne. Det går också bra om informanten vill förhandsgranska eventuella egna citat som förekommer i uppsatsen.

Slutligen vill vi nämna att vi naturligtvis är lyhörda om Ni vill utföra intervjuerna på ett annat sätt än som nämnts ovan.

Med hopp om ett positivt svar!

Med vänliga hälsningar

Jonas Hult & Catharina Grönqvist

Bilaga 3: Arbetsfördelning

Båda författarna har varit involverade i skrivandet av samtliga kapitel i uppsatsen. Detta har skett genom att diskussioner har förts efter studier av relevant material, antingen litteratur eller intervjutranskriptioner (transkriberingarna delades också så att författarna transkriberade samma längd på intervjuerna) under vilka stödord och meningar nedtecknats och sedan renskrivits. Den författare som inte skrev texten korrekturläste sedan denna och gjorde adekvata ändringar. Båda författarna deltog i samtliga intervjuer. Catharina Grönqvist har stått för formgivningen och översättningen av samtliga figurer i uppsatsen.