

**Enligt god sed, inom rimlig tid**  
**Inställningarna till och förutsättningarna för öppen och**  
**delad forskningsdata**

**Agnes Holme**

Examensarbete (30 högskolepoäng) i biblioteks- och informationsvetenskap för  
masterexamen inom ABM-masterprogrammet vid Lunds universitet.

Handledare: Fredrik Åström

År: 2018

## **Title**

In accordance with good research practice, in a timely manner: The attitudes towards and the conditions for open and shared research data

## **Abstract**

With increasing amounts of generated research data, relevant for saving and sharing, as well as higher demands on more open and transparent research, research data management has become an important issue for all involved in the research community. Open access to research data is promoted using arguments concerning public good, research transparency and quality control, the bringing of new research questions, and extended collaboration among researchers.

A large part of the responsibility for the research data management falls on the research or university libraries. Therefore, the aim of this study is to examine the attitude towards and the conditions for open research data on an institutional level – what requirements and guidelines there are, what motives there are for these guidelines, and what stakeholders are involved. The aim is also to compare these requirements and guidelines to those presented by important research funders and academic publishers, to see how compatible they are. This is done by analysing research data management policies from 17 European universities, as well as four publishers of academic journals and four research funders operating in Europe.

The analysis takes on a sociotechnical systems theory perspective and focuses on identifying patterns in the policies' purpose descriptions, their definitions and views of research data, their attitudes towards data sharing and reuse, and their requirements and recommendations regarding open data.

The study showed a generally positive attitude towards open data among the institutional policies, although clear, concrete instructions were lacking. For the most part policies encouraged data sharing, but the decision was left up to the individual researcher or research group. The compatibility with funders' and publishers' data sharing policies appears to be dependent on research discipline, and can vary between different combinations of research institution, funder and journal. However, many institutional policies showed adjustments to external regulations and interests. The universities mostly took on supporting and encouraging roles, but did also show attempts at research governance by presenting incentives for researchers to share their data.

## **Keywords**

library and information studies; LIS; AML; open science; open access; OA; data; research data; research data management; policy; data policy; research data policy; research data management policy; RDM-policy; biblioteks- och informationsvetenskap; ABM; öppen forskning; forskningsdata; forskningsdatahantering; datapolicy; forskningsdatapolicy; forskningsdatahanteringspolicy

## Förord

Inledningsvis vill jag tacka min handledare Fredrik Åström för god vägledning och värdefulla kommentarer. Jag vill även rikta ett varmt tack till Kirsten och Claes-Göran för deras hjälp, uppmuntran och outtröttliga korrekturläsning. Slutligen, tack till Felicia – för att du finns, för att du stöttar, och för att vi har någorlunda samma humor.

# Innehåll

<b>1</b>	<b>Inledning</b>	<b>1</b>
1.1	Problemformulering . . . . .	1
1.2	Syfte och frågeställning . . . . .	2
1.3	Uppsatsens disposition . . . . .	2
1.4	Lista med förkortningar . . . . .	3
<b>2</b>	<b>Bakgrund och tidigare forskning</b>	<b>4</b>
2.1	Open Access . . . . .	4
2.1.1	Open Science . . . . .	5
2.2	Forskningsdata . . . . .	6
2.2.1	Delning av forskningsdata . . . . .	7
2.3	Styrning av forskning . . . . .	10
2.3.1	Datapolicyer . . . . .	11
2.3.2	Aktörer och intressenter . . . . .	12
2.3.3	Främjande av öppen forskning . . . . .	14
<b>3</b>	<b>Teori</b>	<b>17</b>
3.1	Övergripande teoretiska perspektiv . . . . .	17
3.1.1	Sociotekniskt perspektiv . . . . .	17
3.1.2	Systemteori . . . . .	18
3.2	Tillämpning i studien . . . . .	20
<b>4</b>	<b>Metod</b>	<b>21</b>
4.1	Dokumentanalys . . . . .	21
4.2	Material och datainsamling . . . . .	22
4.3	Urval . . . . .	22
4.4	Metod för analys . . . . .	24
4.5	Etiska aspekter . . . . .	25
<b>5</b>	<b>Resultat</b>	<b>26</b>
5.1	Forskningsinstitutioner/universitet . . . . .	26
5.2	Förlag . . . . .	32
5.3	Finansiärer . . . . .	34
<b>6</b>	<b>Analys</b>	<b>38</b>
6.1	Syfte . . . . .	40
6.2	Syner på forskningsdata . . . . .	41
6.3	Återanvändning av data . . . . .	43
6.4	Rekommendationer eller krav . . . . .	46
6.5	Kompatibilitet med förlag och finansiärer . . . . .	48
<b>7</b>	<b>Diskussion</b>	<b>51</b>
7.1	Reflektion över analysen . . . . .	52
7.2	Reflektion över studien . . . . .	54
7.3	Framtida forskning . . . . .	56

# 1 Inledning

Digitaliseringen har fört med sig många nya möjligheter, såväl som nya problem, till forskarvärlden och den vetenskapliga publiceringen och kommunikationen. Det har vuxit fram nya sätt att hantera forskningsmaterial och -resultat, och nya koncept har följt med. Open Access, det vill säga fri tillgång till forskningspublikationer, kan sägas befinna sig i ett skede där det är tämligen väletablerat, men fortfarande mycket omdiskuterat. Det är ett område som utvecklas och är på väg någonstans, och för en del är detta ”någonstans” det mycket bredare begreppet Open Science.

Open Science innebär inte bara öppen tillgång till forskningspublikationer, utan strävar efter att öppna upp alla delar av forskningsprocessen, exempelvis med öppen kod och öppen peer review, och att stärka kommunikationen och delandet av kunskap forskare emellan. Open Science inkluderar även forskningsdata, vilket anknyter till ämnet för denna uppsats. Många argument för Open Access är även tillämpliga i frågan om öppen forskningsdata. Ett av de viktigaste är att alla forskningsresultat som tagits fram genom offentlig finansiering bör göras öppet tillgängliga för att kunna komma allmänheten till gagn. Dessa argument, såväl som en markant ökning i mängden genererad forskningsdata, har gjort forskningsdatahantering till en angelägen fråga. Till detta kommer ökade krav på öppen forskningsdata och fler möjligheter till att dela och återanvända andras data. Ämnet för denna uppsats är därför universitetens forskningsdatahantering, och deras inställning till och förutsättningar för öppen forskningsdata.

## 1.1 Problemformulering

Frågan om öppen tillgång till forskningsdata har blivit allt mer aktuell, och förutsättningarna för det, både de teknologiska och de sociala, förändras snabbt. Allt fler finansörer och förlag, och i förlängningen även forskningsinstitutioner, ställer olika krav på öppen tillgång i någon grad. Detta kan till exempel ses i EU:s satsning på öppen forskning och öppen data i forskningsprogrammet Horizon 2020 (European Commission 2018), eller framväxandet av allt fler OA-tidskrifter och arkiv för öppen publicering av artiklar och data.

Det är dock svårt att få en samlad bild av hur inställningen till öppna forskningsdata egentligen ser ut hos olika aktörer. En intressant ingång till att undersöka detta är att se hur forskningsinstitutioner arbetar med OA och forskningsdataåteranvändning – om och hur de anpassar sig till förlagens och finansierarnas krav, eller om de själva på något sätt driver frågan om öppna forskningsdata, samt vilka motiveringar som finns till det.

Det har gjorts många studier rörande argumenten för och emot öppna forskningsdata och beskrivningar av hur OA utvecklas. Dock är det glesare bland de studier som beskriver hur den utvecklingen inom OA implementeras vid forskningsinstitutionerna. Det finns exempel på studier av amerikanska universitets forskningsdatapolicier, men det hade varit intressant att genomföra någon liknande studie av svenska eller europeiska lärosäten.

Ämnet är högst relevant för biblioteks- och informationsvetenskapen. Universitetsbiblioteken har ett stort ansvar när det kommer till forskarstöd, och en del av de uppgifterna är att ge råd gällande publicering och citering. Akademisk synlighet är viktigt för forskare idag, och exempelvis bibliometrin och dess citeringsanalyser är av stor vikt för lärosätenas forskningsutvärderingar och kan ha betydelse för exempelvis forskningsmedelsfördelning. Vid många universitet faller det även på bibliotekets ansvar att upplysa och utbilda forskarna i god datahantering, och att komma med lösningar till problem gällande lagring och publicering av data. Att finna vägar till en enhetligare datadelning har även ett värde i sig för biblioteksverksamheten, eftersom det skulle öka interoperabiliteten för forskningsdata mellan olika forskningsdiscipliner och -institutioner.

## **1.2 Syfte och frågeställning**

Det övergripande syftet med studien är att undersöka förutsättningarna för forskningsdatadelning och -citering vid europeiska universitet. Tanken är att göra en kartläggning över vilka krav och riktlinjer som finns uttryckta i forskningsinstitutionernas datahanteringspolicyer, samt att jämföra dessa med de bestämmelser som finns hos de stora förlagen och forskningsfinansiärerna. Detta görs för att få en bild av vilka möjligheter och hinder som finns för datadelning och -citering idag, samt vad lärosätena gör för att ta till vara på, respektive överbrygga dessa. Studien kommer därför även att innehålla en genomgång av argument för och emot öppen forskningsdata och hur dessa kan påverka datadelning i praktiken.

De forskningsfrågor som studien utgår ifrån är:

- Vilka riktlinjer och krav gällande öppen data framställs av lärosäten, forskningsfinansiärer samt förlag, och hur skiljer sig dessa mellan de olika aktörerna?
- Utöver universitet, förlag och finansiärer – vilka andra aktörer inom forskarvärlden har betydelse för datadelning och -citering?
- Vilka motiveringar finns för dessa krav och riktlinjer, och vad säger det om de inblandade intressenterna?

## **1.3 Uppsatsens disposition**

De närmast följande två kapitlen hör ihop, och tar tillsammans upp begrepp och fenomen som är nödvändiga för denna studies genomförande och analys. I kapitel 2 redogörs för koncept relaterade till uppsatsens ämne. Tekniska, social och politiska aspekter tas upp, såväl som tidigare forskning om dessa koncept och dess relation till uppsatsämnet i fråga. Kapitlet är relativt omfattande eftersom det är menat att fungera som en del av den teoretiska bakgrunden till studien. I kapitel 3 presenteras det teoretiska ramverket för studien, vilket i huvudsak består av två delar – socioteknisk teori och systemteori för organisationer. I kapitel 4 ges en redogörelse för metoden för studien, vilken är

dokumentanalys av officiella forskningsdatapolicyer. Därefter följer en beskrivning av materialet och urvalet samt av hur analysen av materialet genomförs. I kapitel 5 presenteras resultaten i form av korta sammanfattningar av policyerna. Kapitlet är främst tänkt att tjäna som referens till efterföljande kapitel, och sammanfatta den information som analysen bygger på. Analysen följer i kapitel 6. Uppsatsen avslutas med en diskussion som utgörs av sammanfattningar och slutsatser av analysen, reflektioner kring den egna studien, samt en diskussion kring eventuell vidare forskning.

## **1.4 Lista med förkortningar**

APC – kostnad för publicering i OA-tidskrift, efter ”Article Processing Charge”

DAS – utlåtande om datatillgänglighet, efter ”Data Availability Statement”

DCC – Digital Curation Center

DMP – datahanteringsplan, efter ”Data Management Plan”

FAIR – Findable, Accessible, Interoperable, Reusable

NWO – nederländska vetenskapsrådet, efter ”Nederlandse Organisatie voor Wetenschappelijk Onderzoek”

OA – Open Access

OD – Open Data

ORD Pilot – Horizon 2020 Open Research Data Pilot

RD – forskningsdata, efter ”Research Data”

RDA – Research Data Alliance

RDM – forskningsdatahantering, efter ”Research Data Management”

SND – Svensk Nationell Datatjänst

## 2 Bakgrund och tidigare forskning

Följande avsnitt ger en introduktion till och överblick över de koncept som är viktiga att känna till inför en studie av datadelningens förutsättningar. Bakgrund och information presenteras tillsammans med aktuell forskning inom de olika ämnesområdena. Avsnittet är även tänkt att fungera som en del av den teoretiska grunden för studien. Översikten beskriver allmänt öppen tillgång och forskningsdata, samt delning och citering av denna, för att sedan gå över i styrning av forskning genom policyer och intressenter.

### 2.1 Open Access

Open access holds the promise of moving knowledge from the closed cloisters of privileged, well-endowed university campuses to institutions worldwide.

(Willinsky 2006, s. 33)

Samlingsbegreppet för öppna publiceringsalternativ kallas för Open Access (OA). Peter Suber (2012) beskriver det som att OA-material är digitalt, webbaserat, avgiftsfritt och även befriat från de flesta upphovsrättsrestriktionerna. Enklare uttryckt så innebär OA att barriärerna som hindrar tillgång till materialet har tagits bort, och dessa barriärers karaktär specificerar vilken sorts OA det är frågan om (Suber 2012). OA kan tillämpas i flera olika varianter och på flera olika sorters material, men det finns några mer eller mindre allmänt vedertagna definitioner av OA. En samling av dessa definitioner kallas för BBB-definitionerna, vilka är en kombination av definitioner hämtade från resultaten av tre internationella möten på temat OA, nämligen *Budapest Open Access Initiative 2002*, *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities 2003* och *Bethesda Statement on Open Access Publishing 2003* (Morrison 2009). Utifrån dessa definitioner kan OA kategoriseras bland annat utifrån graden av öppenhet. Detta ger två olika typer av OA – gratis, vilket innebär att materialet finns att tillgå kostnadsfritt, och libre, vilket utöver kostnadsfrihet även innebär att det är fritt att återanvända, det vill säga att något av upphovsrättsrestriktionerna kring materialet har lättats. OA kan även kategoriseras utifrån hur materialet förmedlas, alltså hur OA tar sig uttryck i praktiken. I det hänseendet talas det om två olika sorters OA – grön och guld. Grön OA innebär att materialet publiceras i ett öppet arkiv (Morrison 2009). Många förlag tillåter parallellpublicering av artiklarna som publiceras i deras tidskrifter, oftast i form av preprints, det vill säga förhandstryck. Inom vissa forskningsområden är sådan parallellpublicering mycket vanlig (Suber 2012). Den andra sortens OA kallas guld, och har att göra med OA-tidskrifter, där materialet blir fritt tillgängligt så fort det publiceras. Det finns rena OA-tidskrifter, men även hybrid-tidskrifter där en del artiklar är öppna och andra är avgiftsbelagda (Morrison 2009). En fördel med grön OA är att den, till skillnad från guld OA, kan innefatta andra material än de traditionella forskningsartiklarna, till exempel datamängder (Suber 2012).



### 2.1.1 Open Science

*... a commitment to the value and quality of research carries with it a responsibility to extend the circulation of this work as far as possible, and ideally to all who are interested in it and all who might profit by it...*

(Willinsky 2006, s. 5)

OA kan egentligen ses som en del i ett mycket större system som inbegriper andra delar av forskningsprocessen och forskningsvärlden än bara det publicerade materialet. Bartling och Friesike (2014) skriver om hur forskning egentligen kan ses som kunskapsskapande och -spridning. Historiskt sett har utvecklingen av forskningen till stor del berott på hur detta skapande och denna spridning har gått till. Till exempel har det moderna systemet med publiceringar i vetenskapliga tidskrifter revolutionerat forskningen, men det har även satt en del begränsningar, både tekniska ekonomiska såväl som sociala, för vilket material som kan publiceras. Bartling och Friesike (2014) menar att vi står på gränsen till en ny forskningsrevolution, med nya möjligheter att tillgängliggöra andra former av forskningsmaterial – data, resultat, idéer, och annat – för att på så sätt öka utbytet av kunskap forskare emellan. Open Science är en term som har börjat användas mer och mer under de senaste åren. Detta begrepp är bredare än Open Access och syftar snarare på en öppen vetenskapskultur där resultat delas snabbt och vitt (Bartling & Friesike 2014). Det finns de som hävdar att även termen Open Science är för snäv, eftersom den mest riktar sig till den traditionella synen på vetenskap. Det finns till exempel en snedfördelning i i vilken utsträckning teknologi används i forskningen inom naturvetenskapen, samhällsvetenskapen och de humanistiska vetenskaperna. För att ta hänsyn till de olika vetenskapskulturer som finns, och uppmuntra kommunikation dem emellan, föreslås i stället begreppet Open Knowledge (Sidler 2014).

Oavsett vad fenomenet kallas så finns det ett antal faktorer i dagens forskarvärld som driver på utvecklingen mot mer öppen forskning. Fecher och Friesike (2014) vidareutvecklar detta i sin text om Open Science och dess olika läror. De menar att Open Science som koncept vilar på fem olika läror, som i sin tur grundas i olika motiv och ideologier och styrs av olika mål och intresser. Läror som Fecher och Friesike (2014) tar upp är den allmänna, den demokratiska, den pragmatiska, infrastruktur- och mätningläran.

Den allmänna läran grundas på uppfattningen att forskning bör göras öppet tillgänglig för allmänheten, och arbetar således med att öka tillgängligheten och förståelsen för forskningsmaterial. Tanken är att allmänheten nu kan, och bör, vara mer involverad i forskningen. Exempel på detta är olika medborgarforskningsprojekt, vilka dock även har fått en del kritik – ironiskt nog både för att de blandar in amatörer i seriösa forskningsprojekt, och för att de inte involverar den deltagande allmänheten tillräckligt. En annan aspekt av en allmänna läran är folkfinansiering, eller *crowd funding*, av forskningsprojekt. För att öka allmänhetens insyn i och förståelse av forskningen främjas också nya former av forskningskommunikation, till exempel genom sociala medier och bloggar (Fecher & Friesike 2014).

Den demokratiska läran har att göra med den grundläggande tillgängligheten till forskningsresultat, och grundas på att alla har lika rätt att tillgå forskningsresultaten – särskilt när forskningen är skattefinansierad (Fecher & Friesike 2014). Konceptet syftar på olika sorters material, men fokus ligger på forskningspublikationer och forskningsdata, vilka beskrivs närmare i början av detta avsnitt, samt i nästa.

Den pragmatiska läran ser Open Science som ett sätt att göra forskningen – dess kunskapskapande och -spridning – mer effektiv. Den arbetar för att öka samarbetet forskare emellan och delandet av forskningsmaterial, verktyg och expertis i allmänhet. Den främjar koncept som öppen data och öppen kod (Fecher & Friesike 2014). Nära kopplat till den pragmatiska läran är infrastrukturläran, som ser på Open Science ur ett teknologiskt perspektiv, och fokuserar på de digitala nätverk och verktyg som underlättar forskningen. Forskningsinfrastrukturen är på sätt och vis det som möjliggör de andra delarna av Open Science, eftersom den lägger grunden för de digitala förutsättningar som finns för att dela material över huvud taget (Fecher & Friesike 2014). Mätningläran har att göra med hur vetenskapligt genomslag kan mätas eller utvärderas, och hur detta förändras med utvecklingen inom Open Science. Andra former av utvärderingsmetoder än citeringsanalyser utforskas, till exempel altmetrik, och även andra former av granskning, till exempel öppen peer review.

## 2.2 Forskningsdata

If publications are the stars and planets of the scientific universe, data are the "dark matter" – influential but largely unobserved in our mapping process.

(Brase et al. 2014, s. 183)

Liknelsen som Brase et al. (2014) gör är träffande. Forskningsdatan utgör själva grunden för de vetenskapliga publikationer som läses och refereras till världen över, men historiskt sett har de inte fått samma uppmärksamhet som annat vetenskapligt material. Mycket av det kan bero på svårigheterna med att hantera, definiera och mäta data, liksom fallet även är med mörk materia. Under senare tid har emellertid intresset för forskningsdata ökat, och så även behovet av att utveckla hanteringen av dessa.

Denna studie fokuserar på en gren inom OA-landskapet som har blivit alltmer angelägen, nämligen frågan om öppen eller delad forskningsdata. Något av det första som behöver klarläggas inför en sådan studie är vad forskningsdata egentligen är – en fråga som är mer svårbesvarad än den verkar. Europeiska Kommissionen (2017) definierar forskningsdata som information i form av fakta eller siffror som samlats in för att tjäna som underlag för resonemang, diskussioner och beräkningar. EU-kommissionen syftar främst på data som är digital.

Leonelli (2015) diskuterar olika definitioner av forskningsdata och dess betydelse för exempelvis delbarheten, genom att studera frågan ur ett forskningsfilosofiskt och sociotekniskt perspektiv. Hon noterar att beskrivningar av data främst har fokuserat på den stationära produktionen och tolkningen av data, och att de därför förbiser datans roll

i utbytet mellan människor. Hon menar att data är kommunikationsverktyg för överförande av kunskap och definierar dem som en relationell kategorisering av forskningsresultat som samlats in för att tillhandahålla bevis för specifika kunskapsanspråk. Vidare menar hon även att data inte bara är samlingar av olika värden, utan att de fysiska egenskaperna hos datan är nog så viktiga för förståelsen som de konceptuella. Datans informationsinnehåll kan således inte separeras från formen (Leonelli 2015).

Leonelli (2015) poängterar en paradox som råder inom vetenskapen när det kommer till data. Data behandlas och betraktas som givna och insamlade, men är egentligen producerade genom olika vetenskapliga processer. Detta gör det svårt att se dessa data som objektiva bevis för teorier. Detta har vidare lett till diskussioner kring hur manipulerade forskningsresultat ska vara för att ses som data. Leonelli (2015) föreslår att fokus i stället läggs på vad forskarna ser som data, och i vilken kontext detta sker. Hon menar att alla objekt som uppkommer ur forskning – fotografier, siffror eller annat, som samlas in för att användas som bevis för kunskapsanspråk är data. Objekt kan då, beroende på vilket syfte de har, fungera som data i vissa kontexter, och i andra inte. Detta står i motsats till andra kontextoberoende definitioner, där data anses ha ett fast informationsinnehåll och på så sätt även fixerar tolkningen av dem. Leonellis (2015) definition innebär i stället att samma dataset kan tolkas olika, och därför användas som bevis för olika teorier. Därför, menar hon, kan data inte definieras allmänt, utan måste sättas i relation till forskningskontexten (Leonelli 2015).

Eftersom rollen som data beror på syftet och användandet som bevis, beror den även på möjligheten att förmedla detta bevis till en annan part. Alltså måste objekten vara portabla för att kunna betraktas som data. Data definieras således av deras portabilitet och möjliga bevisnytta (Leonelli 2015). Följande citat sammanfattar Leonellis huvudkriterier för forskningsdata, där del två knyter an till förevarande studies huvudtema, nämligen delad forskningsdata.

Hence, any object can be considered as datum as long as (1) it is treated as potential evidence for one or more claims about phenomena and (2) it is possible to circulate it among individuals.

(Leonelli 2015, s. 817)

### **2.2.1 Delning av forskningsdata**

Data will remain 'dark matter' in scholarly communication unless they are described, curated, and made discoverable.

(Borgman 2015, s. 241)

Öppen och delad forskningsdata är inte någon ny fråga, utan har varit på tal inom forskarvärlden åtminstone sedan 1980-talet. En anledning till att det har tagit fart på senare tid är att forskningsfinansierare började ställa krav på att resulterande data gjordes öppet tillgänglig. Många talande exempel på detta kan tas från USA. Till exempel införde amerikanska National Institute of Health (NIH) 2003 ett krav på att för alla anslag över \$500 000 måste en datahanteringsplan (DMP) utformas. 2010 kom ännu

hårdare krav på DMP från National Science Foundation (NSF). Alla finansierade projekt skulle ha en tvåsidig DMP som skulle klargöra hur forskningsdatan skulle göras öppet tillgänglig inom rimlig tid, och planen i sig skulle genomgå peer review. Detta initiativ från NSF anses vara en vändpunkt i frågan om delad forskningsdata. Finansiärer i Storbritannien började ställa krav på öppen data under 1990-talet, vilket ledde till att Digital Curation Center (DCC) grundades och började erbjuda stöd i forskningsdatahantering, till exempel genom att utforma DMP-mallar. Material från DCC är till hjälp för forskare och forskningsinstitutioner, även utanför Storbritannien och Europa (Borgman 2012).

Argumenten för delad forskningsdata är många. Borgman (2012) tar upp fyra logiska grundmotiv för delad forskningsdata och går igenom vilka intressenter som stödjer sig på vilka grunder samt vilka fördelar dessa intressenter får genom delad data. Grundmotiven är återskapande och verifiering av resultat, allmänhetens intresse, uppkomst av nya frågor och framsteg inom forskningen. Borgman (2012) ordnar dessa utefter om de är till nytta för forskare och dataproducenter eller för dataanvändare, samt om argumenten för delad data är utifrån forskningens behov eller utifrån allmänhetens. Det första motivet är mycket starkt ur ett forskningsperspektiv; att kunna återskapa och verifiera forskning anses som något eftersträvansvärt. Reproducerbarhet är dock komplicerat av många orsaker, till exempel att så många olika sorters data måste anges, såsom laboratorietrustningsinställningar och omvärldsfaktorer. Verifiering är då lättare i ett forskningsdatasammanhang. Det andra motivet kommer utifrån det tidigare beskrivna argumentet för både OA och Open Science – att resultat från offentlig finansierad forskning bör göras öppen till gagn för allmänheten. Detta motiv är tydligt användarfokuserat och ser till allmänhetens intressen. Det tredje motivet är att om data delas hjälper det till att väcka nya frågor, både genom att nya ögon ser på gammal data, och genom att data från olika källor kombineras. Även detta motiv är användarfokuserat, men är så till nytta för både allmänheten och forskningen i sig. Det fjärde motivet grundar sig i att delad data gör forskningen mer effektiv och möjliggör fler framsteg. Det är främst forskningscentrerat, men är intressant för både producenter och användare av forskningsdata (Borgman 2012).

Trots att det finns många argument för öppen data, och att det finns många som är positivt inställda till det, dyker det upp en hel del hinder och problem efter vägen – både individuella och mer övergripande. En betydande orsak som bromsar upp spridningen av öppen data är forskares ovilja till att dela sin data. Detta skriver Borgman (2012) om, och tar då upp att data har större chans att bli delad om delandet i sig på något sätt gynnar den som genererat datan. Att hantera data och göra den öppet tillgänglig kostar både tid och pengar, och datan måste således ha en tänkbar nytta för en lång tid framåt för att det ska anses värt att investera i den. Forskare ställer även detta mot risken för att inte få akademiskt erkännande för den sammanställda datan. Dessutom finns en oro för att datan ska misstolkas eller användas till något som upphovspersonen inte ställer sig bakom. Utöver detta tillkommer oklarheter i frågan om upphovsrätt till data. Känslig data är en stor fråga för sig, i det att det innebär mer jobb med avgöra vilka delar av datan som får göras öppet tillgängliga, samt med att begränsa tillgången de andra delarna.

Frågan om akademiskt erkännande i samband med öppen data har det forskats en del på. Till exempel undersöker Drachen et al. (2016) i en studie från 2016 hur datadelning påverkar citeringsfrekvensen hos artiklar publicerade inom ämnet astrofysik. I studien undersökte de antalet citeringar till artiklar registrerade i NASAs abstractdatabas *Astrophysics Data System* (ADS), och jämförde resultaten för artiklar som hade länkad forskningsdata med de utan data genom bibliometriska citeringsanalyser. Studien visade att artiklarna med länkad data oftare blev citerad än de utan data, men även att mängden artiklar med länkad data var mindre än mängden utan.

Forskarnas inställning är endast en del av det system som påverkar utvecklingen av öppen data. Finansiärerna och den teknologiska infrastrukturen är också viktiga aspekter. Data med statlig finansiering är till exempel mer troliga att delas än sådana med privat finansiering, såsom inom industrin. Data som bär med sig ett ekonomiskt värde för innehavaren delas mer sällan än andra data. Data är även mer benägen att bli delad om proceduren kring delandet är automatiserade och fria från problem med upphovsrättslicenser, samt om datan i fråga hör ihop med en vetenskaplig publikation (Borgman 2012).

Det finns således flera hinder att överkomma för att kunna främja öppen data, men värt att belysa är att en sådan utveckling även kan medföra nya dilemman. Borgman (2012) påpekar till exempel att den stora mängd av olika typer av data som finns ofta leder till någon form av kategorisering. Sådana försök till kategorisering kan dock stjälpa såväl som hjälpa med definitionen av data, eftersom de tenderar att osynliggöra sådana data som hamnar utanför kategorierna. Många kategoriseringar utgår även från de naturvetenskapliga disciplinerna, vars mall inte passar lika bra för samhällsvetenskapliga eller humanistiska discipliner. Wessels et al. (2014) skriver om hur OA till forskningsdata är ojämnt fördelat över forskarvärlden, och hur det i sin tur skapar dataklyftor. Med dataklyftor menas separeringen av forskningsdatan från den ursprungliga publikationen, vilken har uppkommit i och med ökad OA. Datan var tidigare kopplade till en publikation, men vid OA utgör de mer enskilda entiteter som kan användas i helt andra sammanhang, vilket medför en risk för att förståelsen för datan går förlorad.

Buneman, Davidson och Frew (2016) belyser en del av de rent tekniska problem som uppkommer med datacitering. De menar att den generella strukturen på databaserna sätter käppar i hjulet för utvecklingen. Datacitering är invecklat som det är redan nu, med stor variation bland typer av data som citeras, format på citeringarna, samt krav och önskemål rörande vilken information som skall inkluderas i dem. Emellertid finns det olika standarder och rekommendationer för citering tillgängliga för forskarna. Buneman, Davidson och Frew (2016) menar att detta inte räcker. Vanligtvis skriver forskarna inte sina citeringar själva, utan importerar dem från en databas. Detta betyder att, för att sprida användningen av datacitering måste forskningsdatan inte bara utrustas med beständiga identifierare; de måste även tilldelas färdiga, sökbara referenser, liksom andra vetenskapliga publikationer.

För att försöka komma till rätta med de krav som ställs och de problem som uppstår i

anslutning till öppen data utvecklas olika koncept och riktlinjer för att hjälpa till i hanteringen av forskningsdata. Ett exempel på detta är FAIR-principerna, som är till för att göra forskningsdata bättre anpassad för delning och återanvändning – både för människor och maskiner. FAIR står för Findable, Accessible, Interoperable och Reusable, vilket kan översättas<sup>1</sup> med att data ska vara hittbara, tillgängliga, interoperabla och möjliga att återanvända. Dessa grundprinciper kommer med en rad specifikationer som ska fungera som vägledning i den teknologiska delen av datahanteringen (Wilkinson et al. 2016).

## 2.3 Styrning av forskning

And one had to know the initial and final states to meet that goal; one needed knowledge of the effects before the causes could be initiated.

(Chiang 2016, s. 130)

Whitley (2011) skriver om styrning, eller *governance*, som ett mycket brett begrepp, som innefattar alla de processer som används av privata och statliga organisationer för att exempelvis fastställa mål, driva igenom policyer eller skapa medhåll inom systemet i fråga. Processerna kan vara formella eller informella, men har som syfte att styra verksamheten åt något håll. Whitley (2011) använder sig vidare av detta synsätt för att analysera skiftningar i styrningen inom högskolesystem och vetenskaplig forskning, i synnerhet vid statliga forskningsinstitutioner.

All styrning bygger på att den styrande instansen innehar någon sorts auktoritet inom området i fråga. När det kommer till forskningsstyrning kan det handla om exempelvis statlig styrning genom finansiering eller påverkan från andra auktoritära organisationer. Exempel på sådana grupper är stat, forskningsfinansiärer, vetenskapligt ledande forskningsinstitutioner, institutionsstyrelser, forskargrupper och individuella forskare, och Whitley (2011) menar att det är viktigt att uppmärksamma hur dessa grupper kan påverka både valet av forskningsinriktning och hur resultaten sedan utvärderas. Whitley (2011) gör i detta hänseende en träffande beskrivning av forskning som evolutionära kunskapsgenererande system.

Whitley (2011) beskriver vidare hur ökad extern styrning och främjande av specifika forskningsområden, exempelvis genom forskningsfinansieringen, leder till en ökad konkurrens forskningsinstitutioner emellan såväl som ökade behov av marknadsföring av den egna forskningen. Vetenskapligt anseende och erkännande ger materiella belöningar i form av forskningsanslag, vilket ytterligare förstärker prestigehierarkin och konkurrensen inom forskarvärlden. Dessa förändringar har även lett till ökade kvalitetsgranskningar och nya utvärderingsmetoder, för att på så sätt kunna mäta och styra forskningen.

Ett exempel på statlig styrning i en svensk kontext regeringens forskningsproposition från 2016, som inkluderar ett avsnitt om öppen tillgång till forskningsresultat. Där anges att publikationer och forskningsdata som är resultat av offentlig finansiering bör göras

---

<sup>1</sup> författarens översättning

öppet tillgängliga, samt ges förslag på hur och av vem detta ska genomdrivas (Prop. 2016/17:50). Institutionella datapolicyer kan således till viss del ses om en följd av något som är statligt eller externt introducerat, varefter det blir institutionellt adapterat och ger upphov till nya riktlinjer och praktiker.

### **2.3.1 Datapolicyer**

Det är många som skulle gynnas av att det infördes institutionella datahanteringspolicyer (RDM), i alla fall om man ser till vilka aktörer som är drivande i processen. Det går att se på det från två perspektiv – inifrån och utifrån forskningsinstitutionerna i fråga. De främsta pådrivande aktörerna för RDM-policyer, utanför forskningsinstitutionerna, är forskningsfinansiärerna (Rice & Southall 2016). Detta bland annat på grund av ökade krav från OA-rörelsen. Tydligare forskningsdatahantering ökar också effektiviteten, i det att data kan återanvändas och återanalyseras, vilket betyder att finansiärerna inte behöver betala två gånger för att ta fram samma data. Det blir även lättare att förhindra fusk när forskningens resultat går att spåra till de data som finns lagrade (Rice & Southall 2016). RDM-policyer kan även vara en hjälp för forskningsinstitutioner genom att tydliggöra roller och ansvarsområden gällande datahantering, vilket kan underlätta arbetet för att uppfylla andra institutionella policyer, såsom exempelvis forskningsintegritet och informationssäkerhet (Rice & Southall 2016).

Statligt introducerade styrmekanismer för forskning leder ofta till att de i någon form blir institutionellt adapterade. Detta synsätt kan också appliceras på ämnet forskningsdatapolicyer. Ett exempel på detta finns hos DCC och SPARC Europe (2018) och deras arbete med Open Science och, mer specifikt, nationella forskningsdatapolicyer. I en rapport från 2017 analyseras läget för nationella forskningsdatahanteringspolicyer i Europa – hur de är utformade, vem som är drivande bakom dem och hur de ställer sig till Open Science. Utifrån kategorier såsom policytyp, omfattning, koppling till OA och Open Science, krav och rekommendationer, och utveckling av stöd för forskningsdatahantering, kartlägger de förutsättningarna för öppen data på nationell nivå i Europa (DCC & SPARC Europe 2018).

Liknande undersökningar har även gjorts för universitet och forskningsinstitutioner i olika utsträckning. Briney, Goben och Zilinski (2015) skriver om datarelaterade bibliotekstjänster och forskningsdatapolicyer vid amerikanska universitet. De påpekar att studier kring datapolicyer mest har fokuserat på finansiärer, stat eller tidskrifter, och att bibliotekens roll i det hela därför är oklar. I sin studie undersökte de vilka samband det fanns mellan innehållen i policyer som tar upp forskningsdata och vilka datahanteringstjänster som erbjöds vid biblioteket, vid 206 olika universitet. Det visade sig att datahanteringstjänster vid universitetsbibliotek blir allt vanligare; en del bibliotek hade rentav särskilda databibliotekarier. Detta speglades dock inte i policyerna, som ofta var otydliga och förvirrande. Emellertid visade studien även att det fanns ett samband mellan förekomsten av fristående datapolicyer och förekomsten av databibliotekarier och datahanteringstjänster (Briney, Goben & Zilinski 2015). En liknande studie har gjorts av Higman och Pinfield (2015) ur ett brittiskt perspektiv. Deras studie gick ut på en granskning av 37 stycken institutionella forskningsdatahanteringspolicyer hos universitet

i Storbritannien, kombinerat med intervjuer av personer engagerade i utvecklingen av RDM-policyer, vilket analyserades utifrån aktör-nätverk-teori. Resultatet visade på ett komplext nätverk av aktörer både inom och utanför universitetet. Inställningen till delad data var positiv och ämnet främjades i policyerna. Dock förklarades detta ha mycket att göra med inställningen hos de stora forskningsfinansiärerna (Higman & Pinfield 2015). Ett annat sätt att närma sig ämnet presenteras av Young et al. (2017). I sin studie använder de systemteori och organisationsteori för att undersöka hur forskningssystemet möter politisk styrning utifrån, hur det påverkar förändringar i systemet, och hur forskarna själva försöker värja sig mot sådan styrning.

### **2.3.2 Aktörer och intressenter**

Jubb (2013) beskriver vetenskaplig kommunikation idag som ett ekosystem som genomgår konstant förändring, bland annat på grund av de förändrade inställningarna till och förutsättningarna för OA hos olika intressenter inom den vetenskapliga publikationsvärlden. Kommunikationen mellan dessa intressenter är viktigare än någonsin, för att kunna förstå hur dynamiken i forskningsekosystemet fungerar (Jubb 2013).

I en rapport framarbetad inom ett EU-projekt kallat RECODE (Policy RECommendations for Open access to research Data in Europe) skriver Sveinsdottir et al. (2013) bland annat om intressenter i ekosystemet kring öppen forskningsdata. Gruppen av intressenter beskrivs som en brokig blandning av stat, industri, allmänhet och massmedia, och författarna gör ett försök att kategorisera dem enligt deras funktioner i OA-ekosystemet. I huvudsak fem olika funktioner beskrivs: finansiärer och initiativtagare, skapare, distributörer, förvaltare och användare. En studie relaterad till den gjord av Sveinsdottir et al. (2013) gjordes av Ekström (2017), då med fokus på OA-ekosystemet i Sverige och Storbritannien. Utifrån aktör-nätverk-teorin undersökte Ekström (2017) likheter och skillnader i inställningen till OA mellan Storbritannien och Sverige.

#### **Forskare**

Forskarna är de som står i centrum i ekosystemet, och som är i fokus för den vetenskapliga kommunikationen (Jubb 2013). Enligt Sveinsdottir et al. (2013) kan de beskrivas som i huvudsak skapare, men även distributörer och användare av forskningsdata.

De flesta forskare tror att OA gynnar forskningen, men tyvärr är det få av dem som omsätter det i praktiken. Det visar sig att många stöter på hinder i APC-finansieringen, eller bara låter bli att parallellpublicera. Processerna för OA-publicering är fortfarande så komplicerade, och eftersom forskarna ändå har tillgång till en hel del material via sina bibliotek, är det få som lägger ner extra arbete på OA-publicering (Kiley 2013).

#### **Bibliotek**

I och med digitaliseringen har bibliotekets och bibliotekariens roll gått från att vara i huvudsak bevarande och förmedlande till strategiskt utvecklande och främjande



förändring av den vetenskapliga kommunikationen. Biblioteken skapar allt fler samarbeten inom forskarvärlden och jobbar hela tiden med att utveckla forskningsinfrastrukturen och förbättra tillgängligheten och bevarandet (Brown 2013). Bibliotekens mångsidiga roll i forskningsekosystemet kommenteras även av Sveinsdottir et al. (2013), där de tillskrivs huvudfunktionen förvaltare, men även, skapare, distributörer och användare.

Denna bredare roll har lett till att biblioteken även blir mer inblandade i frågor om licenser, affärsmodeller, rättigheter och forskningsutvärdering, vilket knyter biblioteket starkare till forskarvärlden. Det har även lett till att förväntningarna på biblioteket från deras användare har ökat, vilket sätter biblioteket i ett dilemma. Många förväntar sig att allt material ska finnas tillgängligt på något sätt, helst snabbt och enkelt, och biblioteket får i sin tur försöka lösa det, samtidigt som ökade prenumerationsavgifter begränsar från andra sidan av publiceringsvärlden. Därför är OA intressant för just forskningsbibliotek (Brown 2013).

Brown (2013) beskriver forskningsbiblioteken som ett viktigt mellansteg, i det att de jobbar för att ta vara på fördelar från finansierare för sina användare, eller forskare, vilket ökar intresset och engagemanget i öppen forskning, statligt och icke-statligt. Biblioteken testar även gränserna för dagens publiceringsmöjligheter, och undersöker på så sätt hur de kan förbättras och vidgas, och de stödjer institutionen i OA-arbetet genom att tolka och kartlägga OA-landskapet (Brown 2013). Samma resonemang kan appliceras på frågan om forskningsdata. Meningen är inte att biblioteken ska ta hand om forskarnas data, utan de ska samla kunskap om forskningsdataekosystemet och bistå med rådgivning om forskningsdatahantering, till exempel kring hur de olika kraven från finansierare och andra aktörer ska kunna kombineras i praktiken (Brown 2013).

### **Finansiärer**

Forskningsfinansiärer har ett intresse i att resultaten från forskning de bekostat får akademiskt genomslag och blir allmänt kända och tillgängliga, samtidigt som det ska vara ekonomiskt gångbart (Jubb 2013). Frågan om forskningsfinansiering blir därför ofta en diskussion om ekonomisk vinst, även när det kommer till frågor om OA och forskningsdata. Ett exempel är *The human genome project* i vilket investeringar visade sig vara mycket ekonomiskt fördelaktiga, det vill säga att det gav mycket tillbaka från skattemedlen. Detta menar Kiley (2013) visar på de ekonomiska fördelarna OA innebär för forskningsfinansiärer. Kiley (2013) tydliggör detta genom exemplet Wellcome Trust. Wellcome Trust var den första forskningsfinansiären som instiftade en OA-policy, efter att ha påpekat det orimliga i att de själva och forskare inom ämnet inte kunde få tillgång till viktiga framsteg inom malariaforskning, forskning finansierad av Wellcome Trust, utan att betala en extra avgift. Efter det instiftades en policy som deklarerade att alla forskningspublikationer finansierade av Wellcome Trust skulle göras OA, att Wellcome Trust skulle öka anslagen för att täcka OA-publiceringskostnader om nödvändigt, samt att vid beslut om finansiering skulle forskningens syfte och nytta, och inte vilken tidskrift som publicerade den, beaktas i första hand (Kiley 2013). Många forskningsfinansiärer följde i Wellcome Trusts fotspår. Något som är särskilt aktuellt för europeiska forskare idag är EU-kommissionens program Horizon 2020. Det är ett program för forskning och

innovation som sträcker sig över sju år, med en budget på nära 80 miljarder euro. Programmet är tänkt att stärka Europas forskning inom olika områden och göra den mer öppen (European Commission 2018). Alla vetenskapliga publikationer inom programmet måste göras OA. Sedan 2017 har även steg tagits mot öppen forskningsdata genom Horizon 2020 Open Research Data Pilot (ORD Pilot). Denna innebär att forskningsdata genererade inom projekt finansierade genom Horizon 2020 ska göras öppet tillgängliga för återanvändning (European Commission 2017).

Några av fördelarna med OA är ökad synlighet, underlättande av forskningsprocessen och lägre kostnader för vetenskaplig kommunikation (Kiley 2013). Detta kan innebära direkta fördelar för finansierarna, i det att den forskning de bistått i att ta fram får större spridning, materialet som genererades genom forskningen kan användas till annan forskning och på så sätt ytterligare sprida den och sänka kostnaderna för finansierarna, samt sänka kostnaderna för publiceringsprocessen.

### **Förlag**

Enligt Rzepa (2013) är datahanteringen hos tidskrifterna eftersatt. Historiskt sett har forskare haft mycket liten möjlighet att inkludera data i sina publikationer. Detta har delvist lösts genom utvecklandet av digitala datarepositorier, där datan kan lagras sökbar och ofta öppen, med mer informativ metadata än vad som är möjligt i en bilaga till en tidskrift (Rzepa 2013). Den senaste tiden har emellertid mängden data relaterad till publicerade artiklar ökat markant. En del data kan presenteras i eller i anslutning till själva artikeln, men allt mer data presenteras i stället länkad och publiceras och lagras någon annanstans. Detta ställer i sin tur högre krav på utgivarna gällande kvalitetsgranskning, digitala lösningar för tillgång till datan, riktlinjer för data citering samt samarbeten med dataarkiv (Jubb 2013).

### **2.3.3 Främjande av öppen forskning**

Utöver intressenterna beskrivna i föregående avsnitt, finns det en mängd andra organisationer och initiativ som har betydelse i forskningsdataekosystemet. Nedan följer presentationer av ett antal initiativ som på olika sätt verkar för att främja öppen forskning och forskningsdata i Europa.

#### **OpenAIRE**

OpenAIRE 2020-projektet är ett samarbete mellan ett flertal EU-länder, som arbetar för att främja öppen forskning och förbättra synligheten och återanvändandet av forskningspublikationer och -data. OpenAIRE är en sorts utbyggnad av Horizon 2020, och fungerar som en del av infrastrukturen som håller ihop projekten som ingår i programmet. OpenAIRE arbetar bland annat med att ha uppsikt över forskningsresultat ur Horizon 2020-projekt, samt organisera aktiviteter relaterade till dessa, och samordnar tjänster för forskningsdatahantering (OpenAIRE 2017). De erbjuder guider och råd åt forskare, administratörer, finansierare med flera, kring publicering, OA, datahantering och dylikt, och presenterar en sammanställning av öppna forskningsarkiv i Europa.

### **Research Data Alliance**

Research Data Alliance (RDA) är en icke-vinstdrivande organisation, bildad 2013 som ett samarbete mellan EU-kommissionen, det statliga amerikanska National Science Foundation (NSF) och National Institute of Standards and Technology (NIST), och australiensiska Department of Innovation. RDA:s mål är att främja delning av forskningsdata genom att förbättra den sociala och teknologiska forskningsinfrastrukturen (Research Data Alliance 2017). RDA arbetar genom intressegrupper och arbetsgrupper för att utbyta kunskap och upptäckter rörande forskningsdata och för att diskutera problem och möjliga lösningar. Detta resulterar i praktiska rekommendationer för arbetet med att dela forskningsdata (Research Data Alliance 2018).

### **CODATA**

CODATA är kommittén för data vid International Council for Science (ICSU) och arbetar för att förbättra tillgången till och användbarheten av forskningsdata genom globala samarbeten. Forskningsdata ska enligt CODATA vara så öppen som möjligt och så stängd som det är nödvändigt. De strävar efter att förbättra interoperabiliteten och göra datan FAIR. Detta gör de genom att arrangera olika aktiviteter, såsom konferenser och workshops, kring ämnet forskningsdatahantering, både nationellt och internationellt (CODATA 2018).

### **DataCite**

DataCite är en icke-vinstdrivande organisation, vars mål är att förbättra forskarvärldens användning av datacitering. De erbjuder stöd för bland annat forskare, utgivare och forskningsfinansiärer för att öka förståelsen för forskningsdatans betydelse och förbättra förutsättningarna för synlig och delbar data. För att uppnå detta tillhandahåller de till exempel beständiga identifierare i form av DOI:er till samlingar av data (DataCite 2018b). DataCites arbete baseras på samarbete mellan medlemmar från ett flertal olika länder, bestående av bibliotek, statliga myndigheter, datacentra, forskningsuniversitet och andra (DataCite 2018a).

### **Digital Curation Centre**

Digital Curation Centre (DCC) är ett brittiskt centrum för digitalt bevarande, som arbetar med att ge råd åt forskningsorganisationer om lagrandet, hanteringen och delningen av digital forskningsdata. Till exempel erbjuder de guider och checklistor för utformning av datahanteringsplaner och policyer (Digital Curation Centre 2018).

### **Svensk nationell datatjänst**

Svensk nationell datatjänst (SND) är en nationell forskningsinfrastruktur, som drivs av ett konsortium av svenska universitet – Göteborgs universitet, som är värduniversitet och konsortiumledare, Lunds universitet, Umeå universitet, Uppsala universitet, Sveriges lantbruksuniversitet, Stockholms universitet samt Karolinska institutet. Målet är att stödja tillgänglighet, bevarande och återanvändning av forskningsdata. SND erbjuder bland annat datalagring och metadataframställning, och vägledning i olika frågor gällande exempelvis datahanteringsplaner och forskningsjuridik, samt anordnar evenemang och utbildningar rörande forskningsdata (Svensk nationell datatjänst 2018).

### **SHERPA RoMEO och Juliet**

SHERPA, eller Securing a Hybrid Environment for Research Preservation and Access, är en ursprungligen brittisk organisation, grundad 2002 och baserad vid University of Nottingham. Organisationen arbetar för ökad OA genom att stödja utvecklandet av öppna institutionella arkiv. Två av SHERPA's tjänster som är särskilt relevanta för förevarande studie är RoMEO och Juliet (SHERPA 2018). RoMEO är en sammanställning av en mängd policyer kring upphovsrätt och grön OA-publicering för utgivare av forskningstidskrifter. De olika typerna av avtal kodas med olika färger i SHERPA:s databas (SHERPA/RoMEO 2018). Juliet är en databas med information om stora forskningsfinansiärers krav och riktlinjer kring OA-publicering och tillgängliggörande av forskningsdata (SHERPA Juliet 2018). Det finns möjlighet för forskare att samköra sökningar på tidskrifter och finansiärer genom SHERPA:s verktyg FACT (Funders & Authors Compliance Tool), för att kontrollera om de aktuella finansiärernas och utgivarnas OA-policyer är kompatibla med varandra (SHERPA/FACT 2018).

### **ROARMAP**

Registry of Open Access Repository Mandates and Policies (ROARMAP) är en sökbar internationell databas över OA-policyer, utvecklat av University of Southampton. Databasen innehåller information om, och länkar till OA-policyer hos universitet, forskningsinstitutioner och finansiärer, och gör det möjligt att göra detaljerade sökningar bland över 900 olika policyer (ROARMAP 2018).

## 3 Teori

Den teoretiska bakgrunden till studien är en kombination av de begrepp och perspektiv som kan användas för att förklara förekomsten av och innehållet i de datapolicyer som valts ut. Som ramverk för studien läggs ett sociotekniskt perspektiv, det vill säga att analysen utgår ifrån hur teknologiska och sociala förutsättningar påverkar forskningsdatasystemet i stort. Bowen (2009) beskriver en sådan vinkel av dokumentanalysen, där dokumenten ses som sociala fakta, vilka är viktiga att utvärdera i en social kontext. Tillsammans med de rent organisatoriska och teknologiska förutsättningarna för forskningsdatan ger en sådan utvärdering en mer enhetlig bild av datadelningens förutsättningar. Mauthner och Parry (2013) tar upp de sociala förutsättningarna i relation till datadelning, och påpekar att ett problem med dagens datapolicyer är att de främst behandlar problem i själva datadelningsinfrastrukturen, och missar de relationella delarna. De menar att den i huvudsak tekniska synen på datadelning rimmar illa med forskarnas inställning, och att ett mer relationellt synsätt kan hjälpa att främja datadelningen bland forskare (Mauthner & Parry 2013).

Teorin består således av två delar som bygger upp grunden för studien – ett sociotekniskt perspektiv och ett kunskapsorganisatoriskt, där forskningsdatan och dess användning ses som ett ekosystem med både tekniska och sociala komponenter.

### 3.1 Övergripande teoretiska perspektiv

#### 3.1.1 Sociotekniskt perspektiv

Den sociotekniska teorin är tänkt att beskriva samspelet mellan de sociala och de teknologiska komponenterna i ett system, samt deras interaktion med omvärlden. Uppfattningen är att dessa komponenter tillsammans bygger upp systemet och således styr varandra. Inga delar av systemet, sociala eller teknologiska, kan ses som isolerade eller opåverkade av andra delar (Golden 2013).

Det sociotekniska perspektivet definierades på 1950-talet av forskare vid Tavistock Institute of Human Relations i England. Det utvecklades ur en studie av Trist och Bamforth (1951) som inriktade sig på den organisationella strukturen hos kolgruvor. De undersökte vilka sociala och psykologiska följder som långväggsbrytning som metod hade på verksamheten (Sawyer & Jarrahi 2013). Utvecklingen av brytningsmetoden var från början tänkt att effektivisera verksamheten och produktionen, genom att dela upp den och specialisera arbetarna för enskilda uppgifter. Dock visade det sig att denna ordning snarare sänkte effektiviteten, och Trist och Bamforth kom fram till att det berodde på att arbetarnas psykologiska och sociala behov och tillgångar ignorerades. Ur ett rent teknologiskt perspektiv var metoden rimlig, men studien visade att organisationer utgörs av ett samspel av teknologiska och sociala system, varför ett så kallat sociotekniskt perspektiv är mer lämpligt vid förändringar och utvecklingar av dessa verksamheter. Resultaten av studien kom i viss mån att utgöra en motvikt mot tidens massproduktionsperspektiv (Appelbaum 1997).

Teorin har sedan kolgruvestudien utvecklats, och använts flitigt som grund för

organisationsutveckling. Något som betonats då är ökat direkt deltagande av arbetarna i innovativa processer, och ett större inflytande på beslutsfattande. Tanken är att de som själva är involverade i processen eller verksamheten har en bättre insyn i denna och en större förståelse för vad implementerade förändringar får för effekter på andra delar av systemet (Golden 2013). Den sociotekniska teorin har även använts inom informationsvetenskapen, exempelvis av samhälls- och datavetaren Mumford, som arbetade med att tillföra ett mänskligt perspektiv och en social kontext till informationssystemutvecklingen, vilken tidigare hade varit främst datorcentrerad (Sawyer & Jaharri 2013).

Den sociotekniska teorin fokuserar på gemensam optimering av sociala och teknologiska aspekter, där det inte går att rikta in sig på endast en sida, liksom fallet var med kolgruvestudien. Det är ingen mening med att introducera ny teknologi och försöka få människorna att anpassa sig efter den, utan att undersöka hur teknologin påverkar andra delar av systemet, såsom interaktionerna mellan människor (Golden 2013). Det är lätt att tänka sig att vi utvecklar och utformar teknologin för att den ska passa våra behov i ett redan existerande system, men det är även så att introducerandet av ny teknologi i ett system påverkar utvecklingen av människorna, det vill säga de sociala delarna av systemet. Egenskaperna hos och användandet av en teknologi i ett system påverkar personerna som använder den. Teknologin är med och formar hur personerna ser på det egna systemet och den egna rollen och hur de i sin tur förhåller sig till andra personer. Golden (2013) poängterar att teknologi som är tänkt att förbättra ett system, även gör att andra delar av systemet kanske går förlorade, och med det möjliga önskvärda funktioner. Som exempel tar Golden (2013) upp kommunikation genom e-post, vilket visserligen ökar effektiviteten och räckvidden, men även genom sin avskalade form, utan personlig kontakt, ger upphov till större risk för missförstånd. Liknande rubbningar i sociala mönster kan ske vid andra teknologiska förändringar, och det är viktigt att reflektera över vad som går förlorat. En annan viktig aspekt är beroendet mellan de sociala och tekniska delarna av systemet och omgivningen. Förändringar i omgivningen påverkar både de sociala och de tekniska komponenterna i systemet i fråga. Detta får till följd att aktörer med gränsöverskridande roller eller funktioner blir allt viktigare inom organisationer. De observerar och kompenserar för förändringar i systemet och omgivningen så att balansen mellan dem kan upprätthållas (Golden 2013).

### **3.1.2 Systemteori**

Den kunskapsorganisatoriska delen av ramverket för studien utgörs av en del av systemteorin, beskriven bland annat av Vancouver (2013). Systemteori är en bred samling av flera teorier, av vilka de olika organisationsteorierna anses vara del. Tanken med systemteori är att tillhandahålla metoder för att undersöka specifika enheter, eller system, genom att ställa upp riktlinjer för analyser av hur organisationer fungerar – i sig själv och med sin omgivning. De olika systemteorierna kan således ses som verktyg för att utvärdera organisationer. Metoderna för analysen kan vara kvantitativa eller kvalitativa, beroende på vilket system som undersöks och vilken systemteori som appliceras (Vancouver 2013).

Systemteori används i en mängd olika områden för att beskriva olika system, och har sitt ursprung i teorier inom biologi, fysiologi, informationsteori och cybernetik (Stichweh u.å.). Systemteorin introducerades på 1940-talet av biologen Bertalanffy. Bertalanffy observerade att fysikaliska beräkningar och modeller utgick ifrån slutna system, där all information är känd. Sådana system är emellertid inte realistiska i verkligheten. Bertalanffy beskrev det som att alla organismer är öppna system som interagerar med omgivningen. Alla system, stora som små, kan observeras utifrån vad som går in (input) och vad som kommer ut (output). Förändringen beror på processer i undersystem som antingen är kända eller okända för observatören. Poängen är att alla system består av olika nivåer som påverkar varandra, och kunskap om olika nivåer ger olika kunskap om systemet i stort och hur det beter sig. Allmän systemteori går ut på att dessa koncept kan appliceras på alla olika sorters system – fysikaliska, biologiska, sociala, mentala, med mera (Heylighen 1998).

Systemteorin utvecklades vidare av bland annat Ashby till cybernetik, vilket kan beskrivas som studierna av komplexa system utifrån reglerteknik och återkopplingssystem (Heylighen & Joslyn 1992). Detta har även kopplingar till informationsteori. Shannon och Weiner definierade information och överföring av denna som ett urval bland olika alternativ, vilket gör systemteori till en sorts urvalsteori. Detta kan liknas vid modellerna för nervcellers binära förmedlingsfunktion, och vidare till den binära uppbyggnaden av datorer (Stichweh u.å.).

Systemteorin har också spelat en roll inom samhällsvetenskapen, bland annat genom utvecklingen av informations- och kommunikationsteori från cybernetiken. De mest framträdande teoretikerna inom samhällsvetenskaplig systemteori var Parsons och Luhmann, som arbetade med att sätta teorin i en sociologisk kontext. Parsons skrev bland annat om olika typer av sociala system, input/output-analys, samt utbyte och överföring av information mellan system. Luhmann fortsatte på Parsons spår, men med en annan vinkel. I stället för att, som Parson, fokusera på interaktionen mellan olika system, gick Luhmann mer efter Bertalanffys teori kring öppna system, och fokuserade på skillnaden mellan systemet i sig och dess omgivning. Andra viktiga punkter var synen på kommunikation som en selektionsmetod, samt systemens funktion som problemlösare, vilket hör ihop med syftena med systemens uppkomst (Stichweh u.å.).

Det finns som nämnt flera olika systemteorier, beroende på vad de fokuserar på, som kan kopplas till organisationsteori. Rationell systemteori fokuserar på organisationers mål och de kontrollsystem som hjälper till att uppnå och upprätthålla dessa. Naturlig systemteori fokuserar i stället på målen hos delarna, eller individerna, i systemet och interaktionerna dem emellan. När fokus i stället ligger på systemets interaktion med omgivningen handlar det om öppen systemteori (Vancouver 2013). Den sociotekniska systemteorin, utvecklad av Trist, vilken nämns i föregående avsnitt, är en del av den öppna systemteorin.

Idag används systemteori både inom naturvetenskapliga ämnen, i form av allmän systemteori, och i olika former inom samhällsvetenskapliga ämnen (Stichweh u.å.). Den har utvecklats och förgrenats ut i andra teorier, som exempelvis systemdynamik,

agentbaserad modellering och personalhantering för organisationer. Just organisationell systemteori är tämligen väletablerad idag, och merparten av de olika organisationsteorierna bygger på något sätt på systemteori (Vancouver 2013).

Systemteori hör ihop med den sociotekniska teorin i det att den beskriver organisationer som bestående av mindre, interagerande och sammankopplade delar. Den första av två grundläggande uppfattningar som följer av detta är att förändringar i en del av systemet kan ha verkningar i andra delar av systemet – ofta oförutsedda sådana. Den andra är att försök till att införa förändringar i systemet ofta inte leder till någon märkbar förändring alls inom det önskade området. Dessa båda fenomen kan förklaras med att systemet i sig består av flera undersystem, vilka interagerar med varandra, samt att systemets tillstånd vid en tidpunkt ofta är beroende av tillståndet vid tidigare tidpunkter. Trögheten vid förändringar beror på att systemet har egenskaperna av ett så kallat kontrollsystem; det vill säga ett system som genom negativ återkoppling stabiliserar sig själv runt ett specifikt tillstånd. Systemet i sig dämpar störningar i systemet och motverkar således förändringar. Utmaningen blir då att försöka genomföra förändringar utan oönskade bieffekter (Vancouver 2013).

### **3.2 Tillämpning i studien**

Socioteknisk teori och systemteori är båda relativt breda teorier, vilka går att applicera på en mängd olika scenarion. I förevarande studie utnyttjas denna bredd för att genom överlappningen av dessa teorier ringa in ämnet öppen forskningsdata. Teorierna blir således, tillsammans med koncepten beskrivna i kapitel 2, modeller att ha i åtanke vid analysen av primärmaterialet, det vill säga institutionella forskningsdatahanteringspolicyer. Den sociotekniska teorin används, som beskrivet, som ramverk för studien, för att urskilja de teknologiska och sociala aspekter som har betydelse för forskningsdataekosystemet. Att forskningsdatan, med dess användare, användningsområden och tekniska lösningar, kan ses som ett ekosystem bekräftar relevansen av att även använda systemanalysen som grund för studien. Analysen av datapolicyerna, liksom nätverket kring forskningsdatan, består av flera komponenter som tillsammans bildar en helhet. Systemanalysen blir även relevant vid granskningen av förändring och utveckling. Frågan om öppen forskningsdata är aktuell, och det bedrivs arbete inom området på flera håll. Utvecklandet och antagandet av en datapolicy är också något som visar på en vilja att förändra. Frågan är om det förändrar, och i så fall hur och varför, och om de bakomliggande viljorna drar åt samma eller olika håll. Dessa frågor gör teorin om stabilitet och tröghet inom system intressant för en analys av datapolicyer.



## 4 Metod

### 4.1 Dokumentanalys

Den metod som används i studien är dokumentanalys av officiella datapolicyer eller riktlinjer gällande forskningsdatahantering. Dokumentanalys beskrivs bland annat av Bowen (2009) som en systematisk procedur för att utvärdera dokument, vilken kan användas för olika kvalitativa studier.

Dokumentanalys går ut på att hitta, välja ut, tolka och beskriva den data som kan identifieras i dokumenten. Dessa data ordnas sedan efter teman enligt en sorts innehållsanalys. Det är vanligt att dokumentanalysen används i kombination med någon annan analysmetod och annat material, för att kunna validera resultaten genom triangulering. Det gör det lättare att upptäcka och eliminera eller kompensera för eventuell snedfördelning eller partiskhet i studien. Att kombinera kvalitativa och kvantitativa forskningsmetoder kan vara mycket givande (Bowen 2009).

Bowen (2009) beskriver olika anledningar till varför dokumentanalys är en användbar metod, och presenterar argument som att metoden är effektiv, stabil, lättillgänglig, exakt, och omfattande – egenskaper som även synes önskvärda när det kommer till officiella policyer. Dokumentanalys är en bra metod för att undersöka förändring, till exempel då både äldre och nyare, reviderade versioner av samma dokument finns att tillgå för jämförelse. Ett annat sätt är att undersöka arbetsmaterialet som leder fram till det slutgiltiga dokumentet (Bowen 2009). Detta gör det till en lämplig metod för att undersöka policyer och arbetet med att ta fram sådana.

Själva analysen är en iterativ process, med omväxlande läsning och identifierande av teman och organiserande av information i kategorier. Själva analysen i sig används ofta till att formulera nya frågor kring problemet. Det centrala i dokumentanalysen är nämligen vilka frågor som ställs kring materialet. Det är inte bara innehållet som ska studeras, utan själva dokumenten i sig måste också ifrågasättas. En av de viktigaste frågorna är, till exempel, vilket syfte som dokumentet har – för vem det är konstruerat och varför. Även det som utelämnas eller saknas i dokumenten är intressant för analysen (Bowen 2009).

Studier som baseras på dokument kan vara antingen analytiskt fokuserade, då datan i fråga genereras för det specifika forskningsproblemet, eller analytiskt filtrerade, då redan existerande dokument filtreras för att skilja ut de som är relevanta för forskningsproblemet (Gibson & Brown 2011). I denna studie används den senare metoden, varför stor vikt läggs vid specificerandet av material och urval av primärdata. Relevanta frågor att ställa om dokumenten är exempelvis när, varför och av vem dokumentet framställdes, vem den tilltänkta målgruppen är, om det har någon koppling till andra dokument samt om, och i så fall varför, det har gjorts någon betydande revidering av dokumentet (Gibson & Brown 2011).

I dokumentanalysen läggs stor vikt vid vilket material som väljs ut, samt hur och varför (Bowen 2009). Vanligt är att materialet består av flera olika typer av dokument, och att

det då kan vara meningsfullt att sortera dokumenten i olika kategorier (Gibson & Brown 2011). Eftersom materialet för denna studie utgörs av i princip endast en typ av dokument – forskningsdatapolicyer – läggs fokus i stället på att identifiera mönster och teman i dokumentens innehåll som kan dela in policyerna i olika kategorier.

## 4.2 Material och datainsamling

Primärmaterialet för studien utgörs av officiella forskningsdatahanteringspolicyer (RDM-policyer) vid europeiska universitet. Policyerna är antingen specifika RDM-policyer, eller RD- eller OA-policyer som tar upp flera aspekter av forskningsdatahantering och delning och återanvändning av forskningsdata. Det är viktigt för studiens analys att policyerna är så utförliga som möjligt, då det indikerar att de som fastställt policyn har tänkt igenom och tagit ställning till de frågor som väcks kring arbetet med forskningsdata. Policyerna är sådana som finns öppet tillgängliga digitalt, på engelska eller svenska, via universitetens hemsidor.

Originalplanen för denna studie var att undersöka datapolicyer i en svensk kontext, men det visade sig vid uppstarten att inga svenska universitet ännu hade fastställt någon allmän datapolicy. I mars 2018 publicerades emellertid en forskningsdatapolicy för Stockholms universitet, vilken därför inkluderas i studien. Arbetet är även i full gång vid andra svenska universitet med att utveckla forskningsdatapolicyer.

Huvuddelen av studien baseras på de institutionella datapolicyerna, men för att få en mer fullständig bild av forskningsdataekosystemet och förutsättningarna för delad forskningsdata görs även en jämförelse med datapolicyer hos betydande forskningsfinansiärer i Europa och hos stora utgivare av vetenskapliga tidskrifter. Här används RD-policyer eller andra officiella policyer som behandlar publicering, återanvändning och citering av forskningsdata. Policyerna är sådana som finns öppet tillgängliga via hemsidorna.

## 4.3 Urval

Utgångspunkten för urvalet var europeiska universitet, och materialet i fråga måste finnas tillgängligt via universitetets hemsida, vara någorlunda omfattande och vara skrivet på engelska, svenska, norska eller danska. En rapport av DCC och SPARC Europe (2018) som redogjorde för vilka länder som hade nationella datapolicyer gav en fingervisning till var det var bra att börja. Institutionella datapolicyer söktes sedan i dessa länder med hjälp av sökverktyget ROARMAP (2018). Förutom i Storbritannien var dock antalet datapolicyer som uppfyllde kriterierna mycket lågt, varför beslutet togs att insamla merparten av materialet bland brittiska universitet. Beslutet stöts av att Storbritannien är ledande inom arbetet med OA i Europa, och har många initiativ till stöd för delad forskningsdata knutet till sig.

Huvuddelen av urvalet av institutionella datapolicyer görs utifrån en förteckning sammanställd av Horton och DCC (2016). Förteckningen listar drygt 50 olika brittiska universitets datapolicyer, och beskriver deras innehåll genom att markera om, och i så

fall hur specifikt eller strikt, de tar ställning till olika forskningsdatarelaterade frågor, såsom exempelvis definitioner, DMP-rekommendationer, etik och omfattning. Frågorna representeras i en tabell med 13 olika kategorier. För varje kategori kan en policy märkas med en av tre grader - ”inte omnämnt”, ”behandlat, men ej helt specificerat” eller ”fullt specificerat”. För denna studie valdes de policyerna ut som hade flest markeringar i de två senare graderna. Detta för att få så utförliga policyer som möjligt att studera. Ett omnämmande räknas då för lika mycket som en full definition, eftersom det viktiga är huruvida vissa frågor behandlas över huvud taget. Skillnaden i hur institutionerna har valt att förhålla sig till dessa frågor kan dessutom vara intressant för analysen. Urvalsmetoden är således snarare ett sätt att rensa bort de policyer som tar ställning till färre frågor gällande forskningsdata än andra policyer gör. Ingen hänsyn togs här till vilka kategorier det var som utelämnades – endast hur många kategorier. De policyer som valdes ut var de som hade 12 respektive 11 markeringar, vilket uppgick till 15 stycken:

- Buckinghamshire New University
- Cranfield University
- Edge Hill University
- Lancaster University
- London School of Hygiene & Tropical Medicine
- University College London
- University of Bristol
- University of Central Lancashire
- University of Liverpool
- University of Manchester
- University of Newcastle Upon Tyne
- University of Salford
- University of Wales Trinity Saint David
- University of Warwick
- University of York

Storbritannien är ett bra exempel att använda sig av när det kommer till analyser av datahanteringspolicyer, eftersom det finns mycket material att tillgå, och de har kommit långt i arbetet med OA och datahantering i allmänhet. En studie av enbart brittiska datapolicyer hade varit intressant och relevant i sig, men som ett test av representativiteten görs här även en jämförelse med två andra europeiska universitet –

universitetet i Utrecht samt Stockholms universitet.

Fyra forskningsfinansiärer och fyra förlag valdes ut med hjälp av söktjänsterna SHERPA/RO-MEO (2018) och SHERPA Juliet (2018), beskrivna tidigare. De förlag som valdes var de största, med flest tidskrifter, som även har forskningsdatapolicyer digitalt tillgängliga:

- Elsevier
- Springer Nature
- Taylor & Francis
- Wiley

I finansiärskategorin valdes fyra olika organisationer som hade koppling till de länder som representerades i urvalet för institutionella policyer, det vill säga Storbritannien, Nederländerna, Sverige, samt övergripande från EU. De finansiärer som valdes var:

- EU-kommissionen, Horizon 2020
- Wellcome Trust, Storbritannien
- Vetenskapsrådet, Sverige
- Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO), vilket är det Nederländska forskningsrådet

#### **4.4 Metod för analys**

Analysen består av en närläsning av policydokumenten, för att på så sätt identifiera trender och mönster dem emellan. Fokus är i första skedet att undersöka vilka ämnen som tas upp, och hur detta skiljer sig mellan olika policyer. Frågan i denna del är således vilka ämnen som inkluderas och vilka som utelämnas. Detta sammanställs i en tabell, vilken sedan används som utgångspunkt i nästa steg i analysen.

Det anses som givande att kombinera kvalitativa och kvantitativa metoder inom dokumentanalys (Bowen 2009). Denna studie utgörs av en i huvudsak kvalitativ analys av policydokument, men den är av en kvantitativ karaktär, i det att den är tänkt att synliggöra hur vanliga olika ställningstaganden rörande forskningsdatahantering är, utifrån hur frekvent och på vilket sätt dessa omnämns i policyerna.

Steg två av analysen går ut på att utifrån de identifierade kategorierna analysera hur de olika forskningsinstitutionerna hanterar forskningsdatarelaterade frågor. Frågorna utgår från Gibson och Brown (2011), och anpassas för att passa en studie av institutionella datapolicyer. Några huvudpunkter för analysen är:

- Syftet med policyn – vem den riktas till och vem som är avsändaren
- Ämnen – vilka delar som inkluderas eller utelämnas

- Problem – vilka som tas upp, om det görs uttalat eller outtalat, samt om problemen är av teknisk eller social karaktär
- Synen på data – definition, samt om datan behandlas som ett teknologiskt, medialt eller socialt koncept
- Delning – orsaker till att data delas eller inte delas

Utöver analysen av institutionella forskningsdatapolicyer tillkommer en jämförelse med riktlinjerna från förlag och finansiärer. Huvudfrågan är om dessa olika policyer är kompatibla med varandra, om kraven och rekommendationerna från institutioner, förlag och finansiärer arbetar med eller emot varandra, och hur det kan kopplas till relationerna mellan dessa aktörer i forskningsvärlden i stort.

#### **4.5 Etiska aspekter**

Vid forskningsstudier är det viktigt att granska den egna metoden och reflektera över vilka etiska problem som kan uppstå under utförandet eller genom behandlingen av forskningsmaterialet. Ofta gäller det frågor om till exempel behandling av personuppgifter eller anonymisering av intervjupersoner. Eftersom primärmaterialet för denna studie består av offentliga, öppet tillgängliga policydokument är en sådan granskning inte tillämplig i detta fall. Ingen personlig information behandlas och ingen anonymisering är nödvändig. I vissa fall finns enskilda författare för policyerna angivna, men innehållet i policyerna anses ändå uttrycka forskningsinstitutionens ståndpunkt, snarare än individens. Även i de fall då enskild författare anges är policyn granskad och godkänd av någon grupp som representerar institutionen. Detsamma gäller för policyerna från förlag och finansiärer.

## 5 Resultat

### 5.1 Forskningsinstitutioner/universitet

Nedan följer sammanfattningar av var och en av de institutionella forskningsdatahanteringspolicyer som inkluderades i studien. En översikt över dokumenten – policytyp, implementerings- respektive revideringsår, samt var de presenteras – ges i tabell 1.

Tabell 1: Sammanfattning av de institutionella policyerna

Universitet	Policytyp	Utgivare	Antagen	Reviderad
Buckinghamshire New University	RDM	Universitetet	2016	
Cranfield University	RDM	Universitetet	2016	
Edge Hill University	RDM	Universitetet	2017	
Lancaster University	RD	Universitetet	2013	2015
London School of Hygiene & Tropical Medicine	RDM	Universitetet	2014	
University College London	RD	Biblioteket	2013	
University of Bristol	RDM/OD	Universitetet	2015	
University of Central Lancashire	RDM	Universitetet	2015	
University of Liverpool	RDM	Biblioteket	2013	2016, 2017
University of Manchester	RDM	Biblioteket	2012	2015
University of Newcastle upon Tyne	RDM	Universitetet	2017	
University of Salford, Manchester	RDM	Biblioteket	2015	
University of Wales Trinity Saint David	RDM	Biblioteket	2016	2017
University of Warwick	RDM	Universitetet	2011	
University of York	RDM	Universitetet	2014	2017
Universiteit Utrecht	RD	Universitetet	2016	
Stockholms universitet	RD	Biblioteket	2018	

#### Buckinghamshire New University

RDM-policyen för Buckinghamshire New University inleds med ett avsnitt som fastställer universitetets ansvar gällande forskningsdatahantering, vad som menas med forskningsdatahantering, samt vilka statliga direktiv som finns för datahantering. Sedan följer ett avsnitt om policyens omfattning, det vill säga vilka den gäller, samt vad som inkluderas i termen forskningsdata. Som sista, och största del, anges ett antal principer som ligger till grund för universitetets datahanteringsarbete. Det anges att universitetet stödjer öppen data som princip, där det är etiskt, juridiskt och kommersiellt möjligt. Universitetet försäkrar att nödvändig infrastruktur för lagring och delning av data ska finnas tillgänglig, samt att dessa tjänster kontinuerligt ska uppdateras och förbättras. Krav ställs på att forskare ska utforma en datahanteringsplan för forskningsprojekt, samt att de om möjligt ska söka ersättning från finansörer för kostnader för datahantering. Ägandeskap av forskningsdatan ska klargöras innan projektets start, och datan kommer sedan att lagras i minst 10 år från eventuell publicering de har gett upphov till. Till sist

anges vem som är ansvarig för information och stöd kring forskningsdatahantering (Buckinghamshire New University 2016).

### **Cranfield University**

Policyn inleds med en kort redogörelse för policyns syfte, definitionen av forskningsdata, samt vem som har ansvar för forskningsdatahanteringen i ett projekt. Därpå följer ett avsnitt om datahantering. Upprättande av DMP innan ett projekt uppmuntras. Publikationer bör inkludera information om var och hur bakomliggande data finns att tillgå, och digital data bör utrustas med beständiga identifierare. Data ska lagras och finnas tillgängliga i minst 10 år efter publicering. Policyn avslutas med ett avsnitt om hur och var data bör lagras (Stephenson & Everitt 2016).

### **Edge Hill University**

Policyn inleds med ett konstaterande av vikten av en systematisk och effektiv forskningsdatahantering och en optimerad och öppen datainsamling. Därefter följer en definition av forskningsdata samt en beskrivning av policyns omfattning, varpå ett antal principer följer. Huvudansvarig för forskningsdatan i ett projekt fastställs. En DMP måste utformas för projektet, och resulterande data måste registreras vid universitetet. Datan själv lagras i externa arkiv, och endast i undantagsfall i universitetets egna arkiv. Data ska lagras i minst 10 år efter publicering. Universitetet erbjuder råd och vägledning kring forskningsdatahantering. Policyn avslutas med en lista över fördelar med forskningsdatahantering (Edge Hill University 2017).

### **Lancaster University**

Inledningen utgörs av en beskrivning av varför god datahantering är viktigt, samt vad syftet med policyn är. Själva policyn är uppdelad utefter vilka krav och förväntningar som finns på forskare respektive på universitetet. Huvudansvariga för ett projekts datahantering fastställs och olika ansvarsområden beskrivs. En DMP ska skrivas vid forskningsanslagsansökan. All data kopplad till en publikation måste arkiveras och utrustas med tillräcklig metadata, och lagras sedan i minst 10 år. Universitetet bidrar med arkiveringsmöjligheter då lagring i andra nationella arkiv inte är möjligt. Universitetet erbjuder även vägledning och träning i olika datahanteringsrelaterade frågor. Policyn utvärderas årligen. Till sist finns en bilaga med definitioner av forskning, forskare och forskningsdata (Lancaster University 2015).

### **London School of Hygiene and Tropical Medicine**

Policyn inleds med information kring vikten och nyttan av korrekt datahantering, samt vad policyn i sig innehåller. Policyn beskrivs som en del av universitetets strategi för forskningsstyrning, och ansvaret för den ligger på universitetets "Data Management Support Service". Efter det följer ett avsnitt om omfånget, samt vid vilka förutsättningar som policyn kan frångås och vem som beslutar om det. Det anges även vem som har ansvar för datahanteringen i ett forskningsprojekt. På det följer en lista med principer för forskningsdatahantering. Forskaren förväntas följa de krav som finns inom den egna

disciplinen. Ägandeskap och upphovsrätt till datan ska vara tydligt, och helst möjliggöra öppet användande. För varje projekt måste en datahanteringsplan utformas efter rekommendationer från finansiärer eller universitetet själv. All forskningsdata måste registreras hos universitetet och lagras i minst 10 år på ett säkert sätt. Forskare uppmuntras att göra sina data öppet tillgängliga inom rimlig tid. Återanvända data måste citeras, och publikationer bör även inkludera information om var datan finns att tillgå. Kostnader för datahantering bör inkluderas i ansökningarna för forskningsfinansiering. Efter listan med principer följer information om kontakt gällande forskningsdata, samt en lista med definitioner (Knight 2014).

### **University College London**

Inledningen av policyn ger en definition av forskningsdata och en beskrivning datans relevans för forskningen, samt en beskrivning av policyns syfte. Efter det följer ett antal principer. Enligt god forskningssed ska data hanteras så att genomslaget maximeras, och så att datan räknas som primära forskningsresultat. Om data har publicerats eller ska långtidslagras ska den göras så öppen som möjligt. Ansvarsfördelning måste fastställas, så att all genererad data alltid har någon som är ansvarig för dess bevarande. Efter principerna följer en ansvarsfördelning. Studenter ska enligt god forskningssed planera och dokumentera sin datahantering. Även forskare ska planera sin datahantering, i form av en DMP. Fakultetsledningen ska stödja forskarna i deras ansvar över forskningsdatan och utbilda dem i forskningsdatahantering. UCL forskningsdataservice ska erbjuda ett institutionellt forskningsdataarkiv för lagring av data, samt vägledning i forskningsdatahantering. Policyn utvärderas minst vart tredje år (Ayrís 2013).

### **University of Bristol**

Policyn inleds med en beskrivning av betydelsen av forskningsdatahantering, samt syftet med policyn, varpå äganderätten till datan klargörs. Därpå följer policyns omfattning, samt definitioner av forskningsdata och dataförvaltare, vilken har huvudansvar för datan i ett forskningsprojekt. Policyns största del behandlar forskarens ansvar. En DMP bör skrivas innan projektets start, och kostnader för datahanteringen bör inkluderas i forskningsansökan. Forskarna uppmuntras att publicera data på lämpligt, digitalt sätt, och bör registrera sin primärdata hos universitetet. Data som enligt forskaren ligger till grund för en publikation bör lagras i universitetets dataarkiv. Alla publikationer bör innehålla information om hur bakomliggande data finns att tillgå. Universitetets ansvar går ut på att tillhandahålla dataarkiv samt vägledning i forskningsdatahantering (University of Bristol 2013).

### **University of Central Lancashire**

Inledningen beskriver hur universitetet ställer sig till öppen forskningsdata, samt presenterar policyn. Därefter följer en beskrivning av policyns syfte samt vilka den gäller, varpå en lista med principer presenteras. Universitetets ansvar är att tillhandahålla arkiveringsmöjligheter för data och publikationer, samt erbjuda råd och undervisning rörande forskningsdatahantering. Forskare måste följa finansiärernas datapolicier, och kostnaden för datahanteringen bör inkluderas i ansökningen för



forskningsfinansieringen. Det finns krav på DMP vid anslagsansökningar, och uppmuntras vid alla sorters forskningsprojekt. Forskaren måste samarbeta med andra avdelningar på universitetet, såsom biblioteket, för att säkerställa en god datahantering. Huvudansvarig för datahanteringen samt ägandeskap fastställs. Publikationer måste innehålla information om hur bakomliggande data finns att tillgå. Alla data som ska sparas måste lagras i ett arkiv inom 12 månader, och alla data som stödjer publicerade resultat ska så långt det är möjligt göras öppet tillgängliga senast vid publiceringen. I fall då det finns legitima skäl att inte göra datan öppen, exempelvis etiska eller juridiska måste forskaren kunna styrka detta. Datan ska i sådana fall ändå lagras och hanteras på sådant sätt att den kan göras öppen om exempelvis hindren är tidbegränsade (Evans 2015).

### **University of Liverpool**

Först ges information bland annat om när policyn godkändes och utvärderades, samt vem som skrivit respektive godkänt den. Första stycket klargör varför forskningsdatahantering är viktigt, samt vad som är syftet med policyn. Efter det följer ett avsnitt som anger vem och vilken data policyn gäller, samt vad som räknas som forskningsdata. På det följer universitetets datahanteringsprinciper. Ansvar för datahantering samt ägandeskap fastställs. Forskare uppmanas att utforma datahanteringsplaner för forskningsprojekt. Detaljer kring hur datan bör lagras anges. Kostnader för datahantering bör inkluderas i forskningsanslagsansökan. Datan bör lagras minst så länge som finansiären kräver, eller minst 10 år. All forskningsdata bör göras öppet tillgänglig inom rimlig tid, och vara utrustad med lämplig metadata. Publikationer bör innehålla information om var bakomliggande data finns att tillgå (Fairclough 2017).

### **University of Manchester**

Policyn inleds med att fastställa vem som har huvudansvar för datahanteringen i ett projekt. Efter det följer ett antal principer, som bland annat tar upp ägandeskap. DMP är ett krav. Data som troligtvis har ett långsiktigt värde måste bevaras, minst så länge som krävs av finansiärerna. Data förväntas göras öppet tillgänglig inom rimlig tid med få restriktioner. Skaparna av datan kan emellertid beviljas "privilegierad tillgång" under begränsad tid. Publikationer bör innehålla information om hur bakomliggande data finns att tillgå (University of Manchester 2015).

### **University of Newcastle Upon Tyne**

Första delen av dokumentet utgörs här av policyprinciperna som fastställer huvudansvaret för datahanteringen, hur och var data ska lagras, samt vad universitetet åtar sig att bidra med för stöd i datahanteringen. DMP uppmuntras och god metadatastandard i linje med goda data citeringsprinciper för publicerad data är ett krav. Data ska finnas tillgänglig i 10 år efter publicering. Hanteringen av icke-digital data nämns. Det anges även att för data med kommersiellt värde kan en fördröjning för den öppna publiceringen beviljas. Efter huvudprinciperna följer avsnitt som beskriver syftet med policyn och nyttan av god datahantering, policyns omfång och en mer detaljerad beskrivning av ansvarsfördelningen inom datahanteringen. Därpå följer avsnitt om

ägandeskap och licenser, vad som gäller vid delning av data, samt förtydliganden av de inledande principerna (Newcastle University 2017).<sup>2</sup>

### **University of Salford**

Policyn inleds med beskrivningar av syftet och omfånget, samt en lista med definitioner av bland annat forskningsdata. De efterföljande utsagorna specificerar ansvarsområden och ägandeskap, DMP-krav, samt hur och var data bör lagras. Data bör utrustas med lämplig metadata för att förenkla återanvändning. Data måste, när det är lämpligt, göras öppet tillgängliga, men ingen specifik tidsram sätts på detta krav. Datahanteringen ska, om möjligt, bekostas av finansörerna (Baker & Clay 2016).

### **University of Wales Trinity Saint David**

Policyn beskriver utförligt bakgrunden till och nyttan med forskningsdatahantering, och vilket ansvar universitetet har i frågan. Efter det följer en redogörelse för policyns syfte. Del tre redogör för policyns omfång, och definierar då vilka forskningsaktiviteter policyn täcker, vad som betraktas som forskningsdata, vem som räknas som forskare vid universitetet, samt vad som är finansierad forskning. Själva policyn och dess principer presenteras i del fyra. Där anges vem som har ansvar för datahanteringen i ett projekt, vad universitetet bidrar med i form av datahanteringstjänster och vilken standard forskningsdatan ska ha och vilka metadata som ska kopplas till dem. Där finns krav på DMP, före projektets start. Data ska lagras så lång tid som finansörerna kräver. Publikationer bör inkludera information om hur bakomliggande forskningsdata finns att tillgå, och denna data bör utrustas med beständiga identifierare. Data ska göras öppet tillgängliga, men forskargrupper kan beviljas en tids ”privilegierad användning” för att de ska kunna publicera sina resultat. Detta embargo kan vara upp till 12 månader. I policyn nämns även icke digital data, vilka bör finnas tillgängliga vid förfrågan och då digitaliseras. Kostnaden för datahantering bör bekostas av finansörerna. Forskningsdata tillhörande ett akademiskt arbete av en anställd på universitetet tillfaller universitetet. Forskningsdata från en student tillfaller studenten, men ska lagras hos universitetet under lämplig tid. Policyn utvärderas årligen (Briggs 2016).

### **University of Warwick**

Första delen av policyn ger en kontextredogörelse som beskriver bakgrunden till policyn, dess syfte och mål, samt varför god datahantering är viktigt. Därefter följer en definition av forskningsdata. Den avslutande, och största, delen av policyn utgörs av ett antal principer. Data ska hålla hög standard, hanteras enligt gällande riktlinjer, och måste registreras hos universitetet. Äganderätt till data får inte överföras till annan part om inte öppen tillgänglighet kan säkras. Huvudansvarig för forskningsdata i projekt fastställs. En DMP måste skrivas, som minst uppfyller universitetets krav. Data ska lagras i minst 10 år från publicering. Data bör göras tillgänglig för återanvändning under lämpliga förhållanden. Universitetet bidrar med tjänster och vägledning rörande forskningsdatahantering (University of Warwick 2011).

---

<sup>2</sup>Hädanefter används det förkortade namnet Newcastle University vid hänvisningar till University of Newcastle Upon Tyne.

## **University of York**

Introduktionen är en kort beskrivning av syftet med policyn, följt av ett avsnitt som fastställer vem policyn gäller, samt vem som har ansvar för att policyn utvärderas och uppdateras. Principerna för policyn presenteras i ett separat stycke. Där anges detaljer för vilka egenskaper datan ska ha och hur den ska förvaras. DMP är ett krav, och kostnader för RDM bör, om möjligt, inkluderas i forskningsanslagen från finansärer. Om inget annat anges ska data lagras i minst 10 år efter publicering, och den måste registreras i Universitetets CRIS-system. Information om hur relaterad data finns att tillgå måste inkluderas i varje forskningspublikation. Efter listan med principer följer en detaljerad redogörelse för ansvarsfördelningen för RDM vid universitetet. Policyn avslutas med ett längre avsnitt med definitioner och förklaringar av olika former av forskningsdata, samt dess betydelse för forskningen (University of York 2017).

## **Universitetet i Utrecht**

Policyn inleds med en definition av forskningsdata. Därpå följer policyns syfte, omfattning, samt begränsningar. Nästa del beskriver de grundläggande principerna. Data ska lagras i minst 10 år från publicering. All data kopplad till ett forskningsprojekt ska bevaras, i den utsträckning de bidrar till verifieringen av forskningen. Datan måste förseas med tillräcklig metadata. Tillgången till forskningsmaterialet måste vara snabb och användarvänlig. Arkiverad data ska göras öppen för tillgång och återanvändning så långt det är rimligtvis möjligt. Efter principerna följer en beskrivning av ansvarsfördelning. Skaparen av datan har också huvudansvar för den. Forskaren ansvarar för att en DMP utformas vid ett projekts start. Kostnaden för forskningsdatahanteringen ska inkluderas i ansökan för forskningsfinansiering. Föreläsare och handledare ansvarar för att förmedla universitetets riktlinjer gällande forskningsdatahantering till studenterna. Studenterna förväntas sätta sig in i universitetets policyer gällande datahantering. Universitetet ansvarar för att utveckla och implementera en forskningsdatapolicy, samt för att tillhandahålla dataarkiveringsmöjligheter samt erbjuda vägledning rörande forskningsdatahantering (Utrecht University 2016).

## **Stockholms universitet**

Policyn är uppdelad i två delar. Den första delen beskriver de principer gällande forskningsdata som universitetet stödjer. FAIR-principerna, hållbara forskningsinfrastrukturer samt öppen forskning tas upp. Universitetet ålägger sig att tillhandahålla forskarna vägledning och stöd i datahanteringen. Del två beskriver reglerna för forskningsdatan. Offentligt finansierad forskning bör göras öppet tillgänglig så långt det är möjligt. FAIR-principerna ska följas och datan ska lagras på lämpliga plattformar. Handledare ansvarar för att informera studenterna om forskningsdatahanteringspolicyn. DMP är ett krav (Stockholms universitet 2018).

## 5.2 Förlag

Nedan följer sammanfattningar av forskningsdatapolicyerna från de fyra förlag som inkluderades i studien.

### Elsevier

Elseviers policy för forskningsdata finns att läsa på en informationssida som inleder med att beskriva varför data bör delas, vilken roll Elsevier spelar i detta, och vilken sorts data som policyn syftar på. Där presenteras även en rad principer som ligger till grund för datapolicyn. Dessa fastställer att forskningsdata, när det är möjligt, bör göras tillgängliga, kostnadsfritt och med få återanvändningsrestriktioner, samt att forskare bör behålla kontrollen över datan och erhålla erkännande när de delar den. Andra som bidrar till datans värde och återanvändning bör också erkännas. Forskningsdatahanteringen varierar mellan olika discipliner, vilket bör tagas i beaktande när krav ställs på delad data (Elsevier 2018a).

Själva policyn presenteras som en lista över initiativ som förlaget ska följa för att arbeta för och stödja delad forskningsdata. Det handlar om att förenkla processen att publicera data för forskare, förtydliga riktlinjer kring exempelvis DAS, samt guider och verktyg för korrekt dataåteranvändning och -citering. De samarbetar också med att förbättra granskningen och utvärderingen av forskningsdata för att stärka datans roll som självständiga forskningsresultat. Även samarbete med andra initiativ till främjande av delad forskningsdata nämns (Elsevier 2018a).

Policyn innehåller i sig inga krav eller riktlinjer för författare att följa. Dessa finns däremot att läsa om på policyns FAQ-sida. Där framgår att förlaget inte lägger något krav på forskarna att dela sina data, men att det uppmuntras. Emellertid kan det finnas krav från enskilda tidskrifter, och forskarna uppmanas att kontrollera detta själva. Den övergripande policyn, med dess riktlinjer kring stöd för delad forskningsdata, gäller för alla Elseviers tidskrifter, men extra krav kan således tillkomma för enskilda tidskrifter (Elsevier 2018b).

### Springer Nature

Springer Nature har en informationssida om forskningsdatapolicyer och -service, vilken i huvudsak är till för författare. Förlaget presenterar där fyra olika policyer samt information om vilka av förlagets tidskrifter som använder sig av vilken policy. Anledningen till att de har fyra olika förklaras med att olika forskningsgemenskaper har kommit olika långt när det gäller delning av forskningsdata (Springer Nature 2018b).

Policyerna listar krav av olika grad på forskningsdatahanteringen, och kraven byggs på för varje policy. I policy 1 uppmuntras delning av forskningsdata, och en lista över lämpliga databaser för lagring och publicering tillhandahålls. I policy 2 är delning ett krav, så länge sekretess och liknande upprätthålls. Öppet tillgängliga forskningsdataarkiv rekommenderas, och för vissa ämnesområden är de ett krav – huvudsakligen ämnen rörande genetik, där de hanterar stora datamängder, så som DNA-sekvenser. Det

rekommenderas att författaren inkluderar ett Data availability statement (DAS), vilket upplyser läsaren om var och hur datan som studien använder finns att tillgå. Förlaget har sex olika standard-DAS som kan användas. I policy 2 tillkommer även information om den upphovsrättslicens (Creative Commons) som gäller för forskningsdatan. I policy 3 är inkludering av ett DAS ett krav, och det krävs också att all data som finns öppet tillgänglig och har använts i studien också citeras på korrekt sätt. Här tillkommer även riktlinjer för peer review. Personerna som ska utvärdera rekommenderas att kontrollera DAS:et. De har även rätt till att begära ut data som studien bygger på om det behövs för kontrollen. Enligt policy 4 måste all data som studien bygger på göras öppet tillgänglig innan publicering. Det tillkommer även mer detaljerade riktlinjer för peer review (Springer Nature 2018b).

På FAQ-sidan kan författarna finna mer utförlig information om datapolicyerna och forskningsdata, exempelvis vilka forskningsmaterial som omfattas av policyn, vilka tidskrifter de gäller för, när de gäller och när det kan beviljas undantag (Springer Nature 2018a).

## **Taylor & Francis**

Taylor & Francis har en servicesida för författare, där en avdelning informerar om delning av forskningsdata. Där finns ett avsnitt om varför data bör delas, en introduktion till förlagets olika policyer, samt en kort guide till hur data bör citeras (Taylor & Francis 2018c). Mer information om exempelvis vilken data som ska delas och när, beständiga identifierare samt peer review av data finns att tillgå på policyns FAQ-sida (Taylor & Francis 2018a).

Taylor & Francis har fem olika policyer, vilka gäller för förlagets olika tidskrifter. Författare uppmanas att kontrollera vilken policy som används för den aktuella tidskriften vid publicering. De olika policyerna kallas *Basic*, *Share upon Reasonable Request*, *Publicly available*, *Open data* och *Open and fully FAIR*. Kraven byggs på för varje policy. I *Basic* uppmanas författare att dela data tillhörande en publikation, såvida det inte bryter mot några regler kring persondata. DAS, datacitering och beständiga identifierare för data uppmanas. Upphovsrättslicens för datan får väljas av författaren. *Share upon Reasonable Request* innebär att författaren måste göra sin data tillgänglig på rimlig begäran. DAS är ett krav, datacitering och beständiga identifierare uppmanas, och upphovsrättslicens är författarens val. Det är även upp till författaren att avgöra om en begäran kan bedömas rimlig. I *Publicly available* ska författaren göra sina data öppet tillgängliga för allmänheten, men kan göra det under en licens med oklar eller restriktiv återanvändningsbestämmelse. Riktlinjer för DAS, datacitering, beständiga identifierare och licenser är de samma som för *Share upon Reasonable Request*. Med *Open data* måste datan göras fritt tillgänglig, under en licens som tillåter återanvändning, det vill säga CC0, CCBY eller likvärdig. DAS, datacitering och beständiga identifierare är obligatoriska. För den sista policyn, *Open and fully FAIR*, gäller samma regler som för *Open data*. Dessutom måste datan uppfylla FAIR-standard (Taylor & Francis 2018b).

## Wiley

Wileys datapolicy finns att läsa på en informationssida om datadelning och -citering, riktad till författare. Där beskrivs fördelarna med öppen forskningsdata, samt hur förlaget arbetar för att göra forskning, och därför även data, mer öppen. Policyn är uppdelad i en datadelningspolicy och en dataciteringspolicy. Datadelningspolicyn finns i tre varianter, uppdelat efter om den uppmuntrar, förväntar sig eller kräver delning av data. Tidskrifter som använder den första varianten uppmuntrar forskare att lagra data som stödjer en publikations resultat i ett öppet arkiv. Ett DAS bör inkluderas i publikationen med en länk till arkivet. För den andra varianten förväntas forskaren göra datan öppet tillgänglig i ett arkiv, men undantag kan göras för känsliga data, exempelvis persondata. Andra material, så som kod och andra verktyg, bör även göras öppet tillgängliga, så långt det är möjligt. Forskarna förväntas även inkludera ett DAS i publikationen med en länk till arkivet. För den sista varianten gäller detsamma som för den andra, med den skillnaden att öppen data och DAS är ett krav (Wiley 2018).

Dataciteringspolicyn består av en detaljerad redogörelse för hur data ska citeras i publikationer, och den gäller för alla Wileys tidskrifter. Mer information om bland annat vilka data som omfattas av policyn, hur och när den måste delas, peer review av data, samt upphovsrättslicenser finns i policyns FAQ-avsnitt (Wiley 2018).

## 5.3 Finansiärer

Nedan följer sammanfattningar av forskningsdatapolicyerna från de fyra finansiärer som inkluderades i studien.

### EU-kommissionen Horizon 2020

EU-kommissionen har genom Horizon 2020 ORD Pilot tagit initiativ till att främja öppen forskningsdata. Deltagande i ORD Pilot var från början frivilligt, men detta ändrades snart så att öppen forskningsdata var standardinställningen för projekt finansierade genom Horizon 2020. Det är emellertid fortfarande frivilligt i viss mån, då det går att gå ur och avskryta sig ORD Pilot vid giltiga skäl. Mer om det beskrivs nedan. De riktlinjer kring OA-publicering och forskningsdatahantering som deltagarna i forskningsprojekt inom Horizon 2020 måste följa finns att ta del av i forskningsprogrammets OA-guide. Dokumentet behandlar OA till både vetenskapliga publikationer och till forskningsdata, och ger därför en allmän genomgång av OA. En del av denna genomgång tar upp forskningsdata, och beskriver då vad OA till forskningsdata under Horizon 2020 innebär, vad data är, samt varför den bör göras öppen (European Commission 2017).

De krav som deltagarna i Horizon 2020 ORD Pilot förväntas följa är att de måste placera forskningsdatan i ett arkiv – institutionellt eller online – och se till så att de, så långt det är möjligt, finns fritt tillgängliga för andra att tillgå, använda, reproducera och sprida. I anslutning till datan ska det även finnas information om vilka verktyg som behövs för att verifiera resultaten i publikationen, exempelvis programkod och analysprotokoll. Dessa verktyg ska, om det är möjligt, tillhandahållas av projektet. Eventuella kostnader för OA

går bra att räkna in i forskningsanslaget. Tillsammans med kraven för öppen data, ges även råd gällande digitala dataarkiv, beständiga identifierare, upphovsrättslicenser samt forskningsdatahantering (European Commission 2017).

Den data som främst avses i ORD Pilot är bakomliggande data till varje forskningspublikation, det vill säga den data som krävs för att validera resultaten i den aktuella studien, och den ska göras öppen så snart som möjligt. Detta gäller även den metadata som finns för att beskriva datan. Dessutom avses även annan data, exempelvis rådata, som genereras inom projektet, samt samhörande metadata. Denna data tillgängliggörs enligt vad som beslutats och specificerats i projektets DMP (European Commission 2017).

Trots att deltagande i ORD Pilot är standard för deltagande i Horizon 2020, så är det möjligt att lämna den om det finns giltiga skäl. Dessa kan vara till exempel att projektet inte genererar någon data, att delningen skulle bryta mot skyddet av personlig eller konfidentiell data, att delandet bryter mot någon förpliktelse gentemot kommersiella intressen, eller att ett deltagande skulle hindra att projektets mål uppnås. Programmets inställning till forskningsdata sammanfattas med frasen ”så öppen som möjligt, så stängd som nödvändigt”<sup>3</sup> (European Commission 2017).

### **Wellcome Trust**

Wellcome Trust inleder med att beskriva hur de strävar efter att forskningsresultat finansierade av dem kommer till användning för allmänhetens bästa, samt att värdet av dessa resultat optimeras. Efter det följer en beskrivning av varför forskningsdata bör göras öppen. Begränsningar av delningen kan dock vara nödvändig i vissa fall, exempelvis när aspekter som sekretess, integritet eller upphovsrätt kräver det (Wellcome Trust 2017).

Forskarna förväntas göra sin data, programvara och material öppet tillgängliga, med så få restriktioner som möjligt. Underliggande data till publikationer, samt programvara som behövs för att tyda datan ska tillgängliggöras senast vid publicering. Data rörande en folkhälsorelaterad nödsituation måste delas så fort och vitt som möjligt. Vid forskningsansökan måste hanteringen och delandet av resultaten övervägas. Detta redovisas i en resultathanteringsplan (output management plan), vilken utvärderas vid beslutstagande om finansiering. Kostnaden för resultathanteringen betraktas som en del av projektbudgeten, och finansieras således även av Wellcome Trust (Wellcome Trust 2017).

De resultat som avses i policyn är exempelvis dataset, programvara som utvecklats, material såsom antikroppar, eller patent (Wellcome Trust 2018). Resultat bör vara möjliga att hitta samt förses med beständiga identifierare. Användande av andras resultat ska citeras på korrekt sätt (Wellcome Trust 2017).

---

<sup>3</sup>författarens översättning

## Vetenskapsrådet

Finansiering från Vetenskapsrådet kommer med ett allmänt krav på OA-publicering av resultat, men de nämner inte specifikt forskningsdata. Detta finns i stället att läsa om i Vetenskapsrådets förslag till nationella riktlinjer för OA från 2015. Avsnittet om forskningsdata inleds med en bakgrundsbeskrivning, samt en genomgång av argument för öppen forskningsdata. Sedan följer en redogörelse för hur de framtagna riktlinjerna grundas i svensk lagstiftning och hur de påverkas av att forskningsdata betraktas som allmän handling. Det tas även upp hur riktlinjerna förhåller sig till EU-kommissionens rekommendationer kring OA genom Horizon 2020 ORD Pilot. Vetenskapsrådets inställning till frågan är att öppen tillgång till forskningsdata bör vara standard, men att undantag kan göras av exempelvis juridiska, etiska eller kommersiella skäl. Införandet av sådana riktlinjer är emellertid en lång process, varför tillfälliga riktlinjer för åren 2015–2020 föreslås. Riktlinjerna som läggs fram gäller forskningsdata framtagen genom offentlig finansiering (Vetenskapsrådet 2015).

Forskningsdata måste förseas med tillräcklig metadata för att andra forskare ska kunna förstå och återanvända datan. Även information om nödvändig mjukvara, licenser för att använda och sprida datan, samt beständiga identifierare måste inkluderas. Formatet på data och metadata ska följa etablerad standard, och data citering ska ske på korrekt sätt. Vetenskapsrådet betonar vikten av att utveckla nya tekniska lösningar för arkivering och tillgängliggörande av data. En teknisk infrastruktur behöver byggas upp för att systemet med delad data ska fungera. För att driva på utvecklingen och skapa nya incitament för forskare att dela sin data ska ett antal pilotutlysningar med krav på öppen data genomföras. Andra frågor som behandlas i riktlinjerna är nyttan av datapubliceringsplaner vid forskningsansökningar, och utbildningar i datahantering för forskare (Vetenskapsrådet 2015).

I dokumentet beskrivs även en ansvarsfördelning kring arbetet med att tillgängliggöra data. Regeringen bör bland annat ta ställning för öppen forskningsdata, samordna uppdragen i arbetet med tillgängliggörandet av data, och satsa på att utveckla forskningsdatainfrastrukturen. På lärosätenas ansvar ligger att arbeta med och att avsätta medel för arkivering och bevarande av data, samt att öka samarbetet med andra lärosäten i arbetet med att utveckla tekniska system och riktlinjer för delning av data. Vetenskapsrådet ser det som sin uppgift att driva på framtagandet av riktlinjer och successivt införa allmänt krav på öppen forskningsdata. Dessutom ska de aktivt stödja projekt som arbetar med att utveckla forskningsdatainfrastrukturen (Vetenskapsrådet 2015).

Det påpekas att Sverige, genom offentlighetsprincipen, redan har en sorts passiv tillgång till forskningsdata, men att målet med riktlinjerna är ett aktivt tillgängliggörande. Det finns emellertid en del undantag gällande vilken data som inkluderas. Riktlinjerna gäller endast rådata, eller primärdata, och det poängteras i texten att det även kan finnas behov av tillgång till aggregerade data, det vill säga sammanslagna eller bearbetade. Riktlinjerna gäller data som samlats in eller genererats genom offentlig finansiering, och som ligger till grund för en publikation. Riktlinjerna gäller ej data som är insamlade sedan tidigare. Tillgänglighet till data som uppkommer genom samarbeten med



internationella forskare eller aktörer inom industrin måste utredas vidare, eller tas ställning till i de enskilda forskningsansökningarna (Vetenskapsrådet 2015).

Som avslutning beskrivs en del av den oro som finns inför att införa krav på öppen data. Forskare är oroliga för att förlora kontrollen över datan – att incitament för samarbeten ska minska såväl som chanserna för finansiering till ett projekt som innebär omfattande datainsamling. Delad data kan också missförstås eller användas på ett skadligt sätt, och det kan bli lättare för fel att sprida sig med datamängder. En fungerande kvalitetsgranskning av data är därför av stor vikt (Vetenskapsrådet 2015).

### **Nederlandse Organisatie voor Wetenschappelijk Onderzoek**

NWO skriver i sin policy att all forskningsdata som stammar ur statligt finansierad forskning ska följa FAIR-principerna. God forskningsdatahantering hör till god forskning, och är en del av Open Science. Grundprincipen är öppen tillgång till forskningsdata och forskningsdatapolicyn som presenteras omfattar därför all forskningsfinansiering från NWO. Dock ska hänsyn tas till aspekter rörande integritet, säkerhet, etik, upphovsrätt och kommersiella intressen, då begränsad tillgänglighet kan tillåtas (NWO 2018).

Forskare som söker medel hos NWO måste ta ställning till forskningsdatan i två steg. Det första är att en datahanteringssektion måste inkluderas i forskningsansökan. I denna sektion ska forskaren svara på om projektet kommer att generera data, och hur de i så fall kommer att lagras, under och efter projektet. Forskaren beslutar själv och beskriver vilka data som är relevanta att spara. Sektionen ligger inte till grund för beslut om finansiering, men kan användas som utgångspunkt för rådgivning gällande forskningsdatahantering. Efter att finansiering är beviljad ska datahanteringssektionen utvecklas till en DMP, vilken måste sändas in till NWO inom fyra månader för att forskningsmedlen ska betalas ut. DMPn kan justeras under projektets gång. Kostnaden för datahanteringen går bra att inkludera i projektets budget i forskningsansökan (NWO 2018).

NWO ger också en beskrivning av vad som betraktas som forskningsdata. Vilken data som är relevant att spara ska stå i rimlig relation till kostnaden för att hantera och lagra datan. NWO kräver endast att data som är lämplig för återanvändning sparas. Under projektets gång ska datan sparas på ett säkert sätt, och efter avslutat projekt ska den lagras i ett nationellt eller internationellt dataarkiv i första hand, och ett institutionellt i andra hand (NWO 2018).

## 6 Analys

Första delen av analysen går ut på att identifiera relevanta teman i de institutionella policydokumenten, och undersöka vilka datarelaterade frågor som behandlas respektive vilka som utelämnas i dokumenten och vad skillnaden är mellan de olika universiteten. En översikt av dessa teman och hur de omfattas av policyerna presenteras i tabell 2.

De teman som tas upp i tabellen är *Omfattning*, *Definition*, *Delning*, *Undantag*, *Citering*, *DAS* och *Kostnad*. *Omfattning* syftar på beskrivningar av vem policyn gäller, till exempel forskare, forskargrupper, studenter, anställda och så vidare, och i viss mån även vilken forskning eller vilka forskningsområden som berörs av policyn. *Definition* syftar på definitioner av forskningsdata, det vill säga beskrivningar av vilken sorts forskningsmaterial som avses i den aktuella policyn. Med *Delning* menas redogörelser för universitetets inställning till öppen och delad data – om det finns några krav eller rekommendationer på att göra datan öppet tillgänglig. *Undantag* är ett tillägg till förgående tema, i det att den beskriver när universitetets policy för öppen data kan frångås. Det kan vara av exempelvis etiska, kommersiella eller integritetsrelaterade skäl. *Citering* och *DAS* har att göra med återanvändning av data – om policyn tar upp om och hur data får återanvändas och hur det ska ges korrekt erkännande. *DAS* syftar på om policyn nämner något om att information om tillgänglighet till underliggande data ska inkluderas i forskningspublikationer. *Kostnad* syftar på om policyn tar upp hur ökade kostnader för datahantering ska täckas inom forskningsprojekten. I tabellen är längre beskrivningar markerade med ett X, och kortare omnämningen med ett O.

Några teman som är intressanta för analysen är inte inkluderade i tabellen, eftersom de nämns i alla policyer, eller i stort sett alla, eller att de endast tas upp i någon enstaka policy. Dessa teman är *Ansvarsfördelning*, *Arkivering*, *Ägande*, *DMP* och *FAIR*. Med *Ansvarsfördelning* menas om policyn tar upp vem eller vilka som har ansvaret för forskningsdatan och datahanteringen, samt vilket ansvar som instiftarna av policyn, alltså universitetet, tar på sig gällande stöd kring datahantering. *Arkivering* syftar på de regler och rekommendationer som universitetet anger rörande var, hur, när och hur länge forskningsdata ska lagras. *Ägande* syftar på frågor om upphovsrätt till forskningsdata, samt hur och när datan får eller ska byta ägare. Detta ämne har, liksom tidigare beskrivet, inget stort fokus i denna analys. *DMP* syftar på krav eller rekommendationer kring datahanteringsplaner. Med *FAIR* menas om policyn tar upp begreppet FAIR data.

Några tänkvärda observationer finns att göra kring den första överblicken av de representerade temana. Till exempel är *Omfattning* något som, föga förvånande, tas upp i de flesta policyerna. För att en policy ska kunna ha någon verkan bör det rimligtvis anges vem eller vad den gäller. Å andra sidan kan det anses som underförstått, vilket kan vara fallet med Cranfield University och Stockholms universitet. *Definition* och *Kostnad* är två andra teman som är flitigt använda, vilket också är rimligt eftersom de är centrala frågor för datahanteringen. *Delning* – ett tema högst relevant för frågan om öppen data – tas också upp av de flesta policyerna, men i många fall läggs mer fokus på undantagen än på riktlinjerna kring delning av data. Detta kan emellertid än en gång tyda på att inställningen är underförstådd – att öppen data är någorlunda självklar, och att det därför

Tabell 2: Teman identifierade i de institutionella policyerna

Universitet	Ämnen som behandlas									
	Omfattning	Definition	Delning	Undantag	Citering	DAS	Kostnad			
Buckinghamshire New University	X	X		O			X			
Cranfield University						X	X			
Edge Hill University	X	X								
Lancaster University	X	X	X							
London School of Hygiene & Tropical Medicine	X		X	X	X	X	X			
University College London		X	X	X			X			
University of Bristol	X	X				X	X			
University of Central Lancashire	X		X	X			X			
University of Liverpool	X	X	X	X			X			
University of Manchester	O		X	X	X	X				
University of Newcastle Upon Tyne	X		X	X						
University of Salford	X	X	X	X	X		X			
University of Wales Trinity Saint David	X	X		X		X	X			
University of Warwick	O	X	O							
University of York	X	X				X				
Universiteit Utrecht	X	X	X				X			
Stockholms universitet			X							

är undantagen som behöver specificeras. Teman som har att göra med återanvändning av andras data, såsom *Citering* och *DAS*, behandlas i relativt få policyer. Intressant är att det enda universitetet som nämner FAIR-principerna är Stockholms universitet. FAIR är ändå ett aktuellt koncept i Europa, till exempel med tanke på att det tas upp inom riktlinjerna för Horizon 2020 (European Commission 2017). Av någon anledning verkar FAIR-principerna ha fått extra stort genomslag i Sverige. Viktigt att nämna är emellertid att flera av de andra policyerna ställer upp egna riktlinjer för forskningsdata, vilka skulle kunna betyda detsamma som FAIR, men beskrivna med andra ord.

Utifrån de identifierade temana har fem huvudämnen ringats in, för att analyseras i relation till kopplingarna mellan aktörerna i forskningsdataekosystemet och till systemet i sig.

## 6.1 Syfte

Det första ämnet har att göra med syftet med policyerna, och relaterar till den tredje forskningsfrågan till denna studie: *Vilka motiveringar finns för dessa krav och riktlinjer, och vad säger det om de inblandade intressenterna?* Merparten av datapolicyerna har uttalade syften, beskrivna i inledningarna. Ett explicit syfte kan ge en fingervisning om vilket fokus RDM-policyerna har och hur universitetet ställer sig i frågan om öppen data. Emellertid skulle även syften kunna utläsas genom de olika teman som tas upp i dokumenten. Syftet kan därför kopplas till alla teman som beskrivs i föregående avsnitt. Detta avsnitt består av en analys av de övergripande, uttalade syften som kan utläsas i policyerna och de likheter och skillnader som finns mellan dem. En återkoppling görs sedan i diskussionskapitlet.

Många av universiteterna har likartade syftesbeskrivningar i sina policyer. Det går ofta ut på att främja forskningsdata som tillgång, och skydda dess värde genom god datahantering. Sådana formuleringar visar å ena sidan på den nytta och de möjligheter som forskningsdatan tillför forskningen, men de belyser även de ekonomiska aspekterna. Det visar på en kommersialisering av forskningsdatan, och i förlängningen även på ett ekonomiskt beroende av omvärlden.

Andra vanliga syften är att sätta upp ett ramverk för ansvarsfördelningen kring forskningsdatan, och att guida forskare och studenter i datahanteringen. Detta visar på en medvetenhet hos universiteterna kring hur forskningsdatan redan är ett befintligt fenomen som måste hanteras på ett tydligt sätt. På sätt och vis kan sådana formuleringar tolkas som att forskningsdatan eller forskningsdatahanteringen ses som ett problem som behöver lösas för att underlätta vidare forskning. Många av dessa syften beskriver även strävan efter att tillgängliggöra data i större utsträckning, vilket även det visar på en insikt gällande forskningsdatans möjliga värde, men utan den kommersiella biten beskriven tidigare.

En vanlig formulering är att policyerna ska sätta upp principer för datahantering i linje med god forskningssed – ett begrepp som kan betyda lite av varje. Det finns en allmän uppfattning om att tydlighet är en fördel när det kommer till RDM-policyer (Rice &

Southall 2016). Ändå är många policyer otydliga kring flera begrepp, som exempelvis *god forskningssed*. Frågan är om de är medvetet otydliga, och i så fall av vilken anledning.

Något som visserligen behandlas i de flesta policyerna, men endast tas upp i en syftesbeskrivning – den från universitetet i Utrecht – är överensstämmelse med gällande lagar, till exempel när det kommer till personuppgifter eller så kallade känsliga data. Universitetet i Utrecht (2016) skriver uttryckligen att policyn bland annat syftar till att utvärdera forskningsdatahanteringen enligt lagar och förordningar. Det är ett exempel på en förvisso kortfattad men tydlig syftesbeskrivning, till skillnad från de luddiga formuleringarna beskrivna ovan. Liknande exempel på mer konkreta syften finns hos University of Central Lancashire (Evans 2015), Newcastle University (2017), University of Warwick (2011) samt Cranfield University (Stephenson & Everitt 2016). De uttrycker att policyerna ska underlätta uppfyllandet av de krav som ställs på datahanteringen av juridiska och etiska skäl. Newcastle University (2017) anger att syftet bland annat är att skydda känslig data, men även immateriella rättigheter och kommersiella möjligheter. Cranfield University är ännu tydligare, och verkar sätta ord på det som lindas in i syftesbeskrivningarna hos andra universitet. Där anges att syftet är att uppfylla finansierarnas krav på datahantering (Stephenson & Everitt 2016). Att den ekonomiska aspekten är viktig bekräftas här, men värt att notera är att vidare kommersiella intressen i forskningsdatan inte nämns.

## 6.2 Synen på forskningsdata

Detta avsnitt behandlar hur de olika aktörerna ser på forskningsdata – dels hur de faktiskt uttryckligen definieras i dokumenten och vilken sorts data som omfattas av policyerna, men även mer allmänt hur aktörerna förhåller sig till data. Denna del av analysen relaterar mest till temana *Definition*, *Omfattning* och *Kostnad* från tabellen.

En rimlig utgångspunkt i analysen av universitetens syn på forskningsdata är hur definitionerna, ifall det finns några, ser ut i policyerna. Ungefär två tredjedelar av de institutionella policyerna presenterar någon form av definition av forskningsdata eller beskrivning av vilka sorts data som omfattas av policyn. En vanlig formulering är att forskningsdata ”inkluderar, men inte begränsas till” ett antal exempel på olika former av data. Denna form av semidefinition ger möjlighet till öppen tolkning av vad som definieras som forskningsdata, samtidigt som den ger några konkreta förslag. Det ger intrycket av att universitetet vill att policyn ska vara till hjälp och vägledning i datahanteringen, men samtidigt inte begränsa för mycket. Det kan också vara ett sätt att undvika mer komplicerade diskussioner kring vad som är forskningsdata, och göra policyn mer allmän. Mer om detta tas upp senare i detta avsnitt.

Ett annat vanligt exempel är att policyn gäller all data som stödjer en forskningspublikation eller behövs för att förstå resultaten i publikationen. Ibland kompletteras det med en variant som beskriver all data som samlas in och bevaras under projektets gång. Paralleller kan här dras både till EU-kommissionens (2017) och till Leonellis (2015) beskrivning av forskningsdata, vilka tas upp i avsnitt 2.2.

EU-kommissionen (2017) definierar forskningsdata som, främst digital, information som tjänar som underlag för vetenskapliga resonemang. Leonelli (2015) utvecklar detta med att poängtera datans kontextberoende och relationella egenskaper. En del beskrivningar i de institutionella policyerna är mycket lika EU-kommissionens (2017) och Leonellis (2015) definitioner, i det att de anammar den mer filosofiska inställningen till forskningsdata – att de är fakta eller kunskap som underbygger ett vetenskapligt argument eller en teori. University of Bristol (2013) inkluderar till och med datans kommunikativa egenskaper, vilket kan liknas vid Leonellis (2015) resonemang kring forskningsdatans krav på portabilitet.

En del policyer, till exempel den från Lancaster University (2015), går djupare in på svårigheten med att definiera forskningsdata, och resonerar närmare kring den enskilda forskarens ansvar i frågan. Andra, exempelvis Edge Hill University (2017) och University of Warwick (2011), påpekar att det inte går att fastställa en enskild definition av forskningsdata, eftersom den är disciplinberoende, varför en mycket bred definition måste användas. Både diskussionen kring den enskilde forskarens ansvar såväl som den om skillnaderna mellan discipliner kan kopplas till Leonellis (2015) resonemang kring kontextberoende. Data är data för de som tolkar och använder dem som data, vilket innebär att forskningsdata även kan ses som en föränderlig egenskap. Detta gör det svårt att anta en allmän definition av forskningsdata som ska kunna gälla alla discipliner vid olika förutsättningar. En sådan syn på data ger ett intressant tillägg till frågan om policyernas tydlighet, i det att det kan ge en motivering till medvetet otydliga formuleringar. Tydligheten i definitionerna av forskningsdata i de institutionella policyerna varierar – från korta fastställanden som ”all relevant data som genereras eller samlas in” (University of Manchester 2015), via de konkreta exemplen, till de mer filosofiska diskussionerna. En medvetet otydlig formulering kan vara ett sätt att kunna utforma en allmän policy som ska kunna vara applicerbar i ett flertal vitt skilda fall inom forskningen vid universitetet.

Trots en mängd små skiftningar i definition så är beskrivningarna av forskningsdata i de olika policyerna likartade. En punkt där det emellertid uppstår en viss diskrepans är i frågan om digitala och analoga eller icke-digitala data. De flesta policyerna utgår ifrån att forskningsdata består av digitalt material, men inkluderar även analoga data. Det finns emellertid undantag, som exempelvis Newcastle University och University of Wales, vilka har separata avsnitt för analoga data och hur dessa ska hanteras (Newcastle University 2017; Briggs 2016). Detta skulle kunna bero på något disciplinspecifikt – att dessa universitet bedriver forskning som genererar mycket analoga data, och därför måste ta ställning till dessa frågor i högre grad än andra. Det kan också bero på att den mesta av den nyare forskningsdatan är digital, och att datahanteringspolicyerna är nya och kanske inte verkar retroaktivt. Analoga data blir då mer ovanliga undantag i små kvantiteter än en del av den egentliga data som åsyftas i policyn.

Kostnaden för datahanteringen och hur den ska lösas tas upp av majoriteten av policyerna, och det uttrycks på ungefär samma sätt – att kostnaden ska inkluderas i ansökan till finansierarna. Det tyder på en praktisk, kanske realistisk syn på data. God datahantering är inte något som kommer av sig själv, utan det måste tas med i

beräkningarna att det krävs tid, arbete och pengar för att få det att fungera.

Synen på forskningsdata hos universiteten byggs således upp av flera olika komponenter. De ekonomiska aspekterna är framträdande på flera ställen – så även de relationella. Sättet på vilket forskningsdata beskrivs i policyerna kan ge en fingervisning angående huruvida de ses mer som problem eller som möjligheter för forskningen.

Förlagens definitioner av forskningsdata är rätt lika universitetens. Forskningsdata beskrivs som det material som underbygger publikationens resultat. Intressant är att Springer Nature tar upp frågan om huruvida fysiska material, såsom prover och cellkulturer, räknas som forskningsdata. Policyn är att det allmänt inte gör det, men att det för enskilda tidskrifter kan gälla annorlunda, och att en allmän policy för tillgång till material kan komma att behöva utvecklas i framtiden (Springer Nature 2018a). Detta skulle kunna tyda på att medvetenheten och förståelsen kring vad som är forskningsdata utvecklas även inom förlagsvärlden.

När det kommer till finansierarna så är det Horizon 2020 som har den definitionen som är mest lik universitetens och förlagens definitioner av forskningsdata. Forskningsdata syftar där på information som samlas in för att användas som grund för resonemang eller beräkningar (European Commission 2017). Vetenskapsrådets (2015) definition är snarlik, med tillägget att fysiska föremål i sig inte räknas som forskningsdata, men väl digital information om dessa föremål. Här har resonemanget tagit en något annan riktning än hos Springer Nature, där fysiska föremål i vissa fall kunde betraktas som data. Intressant är att det ändå tas upp, men utelämnas i de institutionella policyerna. De två kvarvarande finansierarna, Wellcome Trust och NWO, bryter av mot de andra med sina bredare definitioner av forskningsdata. NWO (2018) gör gällande att data både är material som samlas in eller genereras, och sparat analyserat material. Data kan vara både digitalt och icke-digitalt, och inkluderar då även fysiska material, såsom exempelvis prover. Emellertid är de data som behöver sparas, enligt NWO, endast de som är relevanta för återanvändning, och detta kan beslutas bland forskarna inom forskningsdisciplinen i fråga. Wellcome Trust (2017) kan sägas ha den bredaste definitionen av forskningsdata. Policyn hänvisar mer till forskningsresultat än till data, och inbegriper både dataset och mjukvara såväl som material.

### **6.3 Återanvändning av data**

Detta avsnitt behandlar hur universiteten ställer sig till återanvändning av data, om det nämns över huvud taget. Det har att göra med temana *Citering*, *DAS*, *FAIR* och *Ägande*, till viss del.

I många av de institutionella policyerna finns ett avsnitt som beskriver hur universitetet stödjer koncept som Open Science och delad data, men det är få som visar sig ha några tydliga riktlinjer för hur detta ska uppnås eller eftersträvas. En allmänt positiv inställning till öppen data kan emellertid skönjas i flera policyer, till exempel från Edge Hill University (2017) som inleder med ett rent praktiskt ställningstagande. Enligt den är datainsamling kostsamt och tidskrävande, varför det ska optimeras och tillgängliggöras

för återanvändning där det är möjligt. Här kan paralleller dras till den sociotekniska teorin, som i mycket handlar om optimering och sammanlänkning av tekniska och sociala aspekter av ett system (Golden 2013), eller som här – forskningsdatahanteringen vid ett universitet.

Edge Hill University (2017) presenterar även en del argument för delad data, vilka går väl ihop med Fecher och Friesikes (2014) läror om Open Science, samt Borgmans (2012) modell för motiv för öppen data, beskrivna i kapitel 2. Det första argumentet är att offentligt finansierad data tillhör allmänheten och bör användas för dennes bästa. Paralleller kan här dras till Fecher och Friesikes (2014) demokratiska lära. Nästa argument rör bättre reproducerbarhet och validering av forskning, vilket även Borgman tar upp i sin modell (2012). Den pragmatiska läran (Fecher & Friesike 2014) återfinns i argumentet om ökat samarbete och bättre relationer forskare, universitet och finansiärer emellan. Delad data antas också ge ökat akademiskt anseende för forskare och universitet, vilket går i linje med mätningläran (Fecher & Friesike 2014), samt utveckla forskningsfärdigheter, vilket kopplar till infrastrukturläran (Fecher & Friesike 2014). Trots att dessa beskrivningar inte anger några konkreta riktlinjer, så kan de ändå anses visa på vilken grundinställning till öppen data som ligger till grund hos Edge Hill University.

Liknande inställning återfinns hos University of Central Lancashire som skriver att de vill främja öppen tillgång till forskningsdata för att underlätta återvinning, citering, delning och samarbete, samt stödja en mer öppen och transparent forskningskultur. Forskarna uppmanas att anstränga sig att försäkra en öppen delning av datan (Evans 2015). University of Central Lancashire, tillsammans med University of Salford, är också det universitet som bidrar med en av de mest konkreta riktlinjerna gällande öppen data i de undersökta policyerna. Där anges att all data ska placeras i lämpligt arkiv inom 12 månader från att de samlades in eller genererades, samt att de ska göras öppet tillgängliga senast vid datumet för publikationen de understödjer (Evans 2015; Baker och Clay 2016). Detta är, som sagt, exempel på särskilt explicita riktlinjer – något som det är ont om i flera av de resterande universitetens policyer.

Flera av universiteten uttrycker att data ska göras tillgängligt så öppet som möjligt, ibland med tillägget att data ska tillgängliggöras när det är möjligt, eller under lämpliga förhållanden. En vanligt formulering är att data ska tillgängliggöras enligt god forskningssed, vilket egentligen går att tolka på många olika sätt, beroende på exempelvis vilken data det gäller, vilken forskningsdisciplin det rör sig om, och en rad andra omständigheter. Ett annat flitigt återkommande uttryck är att tillgängliggörandet ska göras inom rimlig tid, vilket, liksom uttrycket om god sed, går att tolka på en rad olika sätt.

Trots att många av policyerna är tämligen lika i sina inställningar till öppen data, så är det några exempel som sticker ut och är intressanta att belysa. Stockholms universitet är den enda som använder sig av FAIR-principerna för att fastställa inställningen till öppen data. Många av de andra policyerna ställer upp egna riktlinjer om hur forskningsdata ska sparas, men ingen uttrycker det i form av FAIR. En förklaring kan vara att FAIR inte har



fått lika stark förankring i Storbritannien som det verkar ha fått i Sverige. En annan, kanske troligare, anledning kan vara att FAIR är ett relativt nytt koncept och många av de brittiska policyerna är några år gamla, till skillnad från Stockholms universitets policy, som är från 2018.

University of Liverpool utmärker sig också på en punkt. I sin policy anger de att initiativ till delad data inom universitetet kan belönas genom olika processer för utmärkelser och erkännande (Fairclough 2017). Detta ger ett specifikt exempel på hur enskilda aktörer, såsom forskare, kan dra nytta av att tillgängliggöra sin data, och hur universitetet konkret försöker påverka läget för öppen data. Exemplet från University of Liverpool visar även på intressanta kopplingar till Whitleys (2011) beskrivning av institutionell styrning och konkurrens inom forskning, i det att det skiftar fokus i policyn från stöd till styrning. I stället för, eller utöver att, ställa upp rekommendationer och krav försöker universitetet skapa incitament för forskarna att dela data genom att erbjuda belöningar. Detta kan liknas vid att ökat akademiskt anseende och genomslag kan leda till ökade forskningsanslag (Whitley 2011). Det visar också på nya utvärderingsmetoder, där forskningen värderas utifrån hur öppen den är. Paralleller kan även dras till Vancouver (2013) och systemteorins beskrivning av försök till förändring inom ett system, och hur målen för de enskilda individerna påverkar den förändringen.

Något som också är relevant i diskussionen kring öppen data, är hur inställningen är på andra sidan själva delandet – på mottagarsidan. Inte många av de institutionella datapolicyerna tar upp dataåteranvändning eller data citering och hur det ska gå till. Det visar på en stor forskningslucka inom ett ämne som är viktigt att utforska om användandet av delad data ska växa. DAS nämns visserligen av flera policyer, men det kan egentligen också räknas till det arbete som utförs för att underlätta för andra att återanvända den egna insamlade eller genererade datan. Data citering och användning av data från tredje part tas upp av exempelvis University of Bristol (2013) och University of Manchester (2015), och särskilt ingående av Newcastle University (2017) och University of Salford (Baker & Clay 2016). Till exempel tas det upp hur viktigt det är att få reda på vilka restriktioner som gäller för den data som ska återanvändas. Kanske får den återanvändas, men inte spridas vidare öppet. Sådana hinder skulle potentiellt kunna skapa problem om det kolliderar med finansiärers krav på öppen tillgång.

Bestämmelserna kring huruvida data ska göras öppet tillgängliga är således relativt otydliga, men undantagen från dessa är desto mer explicita. I grund och botten gäller att lagar, kontrakt och avtal med finansiärer står över alla riktlinjer som anges i policyerna. Efter flera principer läggs det in att de endast gäller så länge de inte bryter mot några upphovsrättsliga, juridiska, etiska eller integritetsrelaterade bestämmelser. Vanligt är även att inkludera kommersiella intressen eller förpliktelser i denna lista. Hos vissa universitet, som exempelvis University of Manchester (2015) och Newcastle University (2017), går det att ansöka om en embargo period för sina data. Detta innebär att forskaren, eller forskargruppen, får så kallad privilegierad användning under en begränsad tid. Under denna tid får de ensamrätt till datan och behöver inte göra dem öppet tillgängliga för andra. Anledningen är att data som ligger till grund för möjliga patent eller kommersiella intressen ska få utnyttjas i första hand av de som tagit fram

dem. Ett av de universitet som har ett sådant embargoalternativ är University of Central Lancashire. De knyter det ändå till öppen tillgång, genom att tillägga att embargo-belagd eller undanhållen data ska lagras och hanteras på sådant sätt att de lätt kan tillgängliggöras ifall omständigheterna förändras (Evans 2015).

## 6.4 Rekommendationer eller krav

Denna del av analysen samlas kring den första forskningsfrågan: *Vilka riktlinjer och krav gällande öppen data framställs av lärosäten, forskningsfinansiärer samt förlag, och hur skiljer sig dessa mellan de olika aktörerna?* En viktig punkt här är just skillnaden mellan rekommendationer och krav hos de institutionella policyerna. Detta relaterar till flera av temana i tabell 2: *Omfattning, Ansvarsfördelning, Arkivering, DMP, Delning, Undantag, FAIR och Ägande*. Tanken här är att undersöka trender i vilka krav som ställs och vilka rekommendationer som ges.

Något tänkvärt som visar sig i analysen av universitetets olika krav och rekommendationer gällande forskningsdata är den varierande graden av tydlighet och specificitet. Detta är genomgående i de flesta temana, och skapar ett mönster av gemensamma grundprinciper men med olika tillägg som specificerar riktlinjerna i policyerna. Som ett första exempel kan tas omfattningen av policyerna, det vill säga vilka de gäller. Den grundläggande ståndpunkten är att policyn gäller all personal, och möjligtvis forskarstudenterna, vid det aktuella universitetet, oavsett extern eller intern finansiering. Vissa universitet specificerar vidare till att omfatta alla som utför forskning vid universitetet, och ytterligare några till alla som deltar i skapandet eller insamlandet av forskningsdata.

Den varierande graden av tydlighet återfinns i riktlinjerna kring lagring eller arkivering av forskningsdata. En vanlig metod är att endast ange att forskningsdata ska lagras i lämpligt format, vilket i sig skulle kunna betyda nästan vad som helst. Ett annat sätt att uttrycka det på är att forskningsdata ska hanteras enligt god forskningssed. En sådan formulering lämnar mycket upp till den enskilde forskaren att besluta om. Andra universitet väljer att specificera formen på den lagrade datan på olika sätt. En av de vanligaste formuleringarna är att datan ska lagras på ett sätt som går i linje med avtalsmässiga, juridiska, etiska och andra regulativa krav. Denna inställning visar tydligt på universitetens och forskningens beroende av externa aktörer, såsom exempelvis finansiärer. Specificeringar som i stället fokuserar mer på själva datan och dess värde är sådana som anger att datan ska lagras säkert, det vill säga på ett sätt som skyddar mot dataförlust och -korruption och mot obehörig användning. Ett tredje exempel på specificering anknyter till datans kommunikativa värde genom delning och öppenhet, genom att ange att data ska lagras på ett sätt som underlättar tillgängligheten.

Vilken data som bör sparas och hur länge är också något som varierar mellan de olika universiteten. Den mest grundläggande formuleringen är att all relevant data bör sparas. Lite mer specifikt blir det i de policyer som i stället fokuserar på all data som underbygger publikationer. En annan vanlig formulering är att all data som anses ha långsiktigt värde ska sparas. Frågan blir då vem eller vad som avgör när data är relevant

eller har ett långsiktigt värde. I stort sett alla policyer fastställer att forskningsdata ska sparas i minst tio år efter publicering. Newcastle University (2017) presenterar en något modifierad version av denna princip, där datan i stället ska sparas i tio år efter senaste nyttjandet. Detta är en intressant variant, eftersom den tyder på ett försök till selektion – ett försök att bevara aktuell data. Att se till hur ofta och hur mycket datan används skulle kunna vara ett sätt att avgöra relevansen för den. University of Wales och universitetet i Utrecht har båda till synes mer specifika riktlinjer kring arkiveringstid, utöver grundprincipen om tio år. Den förstnämnda har lagt till att data ska sparas så länge de fortfarande är av värde för forskaren och forskarvärlden. Den senare formulerar det i stället så att så länge datan är relevant för verifieringen av forskningen den stödjer, så ska den sparas. Båda dessa kan ses som lite mer specifika än andra policyer, men egentligen kan arkiveringstiden tolkas till att kunna vara hur lång eller kort som helst, utefter vad den dataansvarige bestämmer sig för.

Det är många frågor och teman som avhandlas i de olika policyerna, och allt analyseras inte här. Ett par exempel på vissa speciella riktlinjer är emellertid värda att ta upp. Det första är från Newcastle University (2017), i vars policy det tas upp hur mycket datautrymme varje forskningsprojekt får att förfoga över vid universitetet, vilket är 0.5TB. Det återkopplar således till det teknologiska perspektivet, vilket annars verkar tendera att falla bort i de institutionella policyerna. Detta skulle kunna bero på ett teknologiska aspekterna anses underförstådda. Många universitet kanske inte heller har koll på hur mycket lagringskapacitet som finns på universitetet, eller hur mycket som behövs. Bra att ha i åtanke här är den sociotekniska teorin – att tekniken inom systemet påverkar människorna och praktiken likväl som tvärt om (Golden 2013). Exemplet från Newcastle University visar på hur teknologin har anpassats efter forskarnas och universitetets behov, men kan även ses som ett exempel på hur teknologin kan användas för att påverka andra delar av systemet, genom att informera om möjligheter och väcka frågor.

Ett annat exempel på riktlinjer är från University of Central Lancashire där samarbetet mellan forskare, eller datahanterare, och biblioteks- och informationservice lyfts fram och uppmuntras (Evans 2015). Biblioteken är annars relativt osynliga i policyerna, även hos de universitet där policyerna faktiskt presenteras via bibliotekens hemsidor. Än en gång kan det bero på att bibliotekens medverkan i datahanteringen är underförstådd, men det kan också tyda på att bibliotekens roll ännu inte är klarlagd, vilket kan liknas vid Browns (2013) beskrivning av den senaste tidens utökning av bibliotekets uppgifter.

Som en avslutning på granskningen av kraven och rekommendationerna från universiteten passar några allmänna kommentarer om policyernas formuleringar in bra. Trots att detta inte är en språklig analys, så ger det ett intressant perspektiv till analysen. Skillnaden på krav och rekommendationer är, liksom många av de olika typerna av riktlinjer, uttryckt varierande tydligt i de olika policyerna. Ofta utgörs texten av tvetydiga begrepp som kan tolkas olika beroende på situation, vilket visar på att inte bara datan i sig är kontextberoende, utan även policyerna. När ord som *Should* dyker upp i policyernas principer, ger det dem karaktären av riktlinjer i stället för något som ska eller behöver följas. Formuleringar med ord som *Must* eller *Will* framstår som starkare, vilket

ställer högre krav på forskarna som policyn rör. Frågan blir då vilken metod som fungerar bäst för att förmedla och driva igenom policyns syften och mål.

## 6.5 Kompatibilitet med förlag och finansiärer

Denna del består till viss del av en övergripande jämförelse mellan de olika förlagens policyer, respektive mellan de olika finansiärernas. Fokus ligger emellertid på hur dessa policyer kan kombineras med de institutionella policyerna och vilka problem som kan uppstå dem emellan.

De fyra förlagen som inkluderats i studien visar prov på olika angreppssätt gällande kraven kring öppen forskningsdata. Mest lika varandra är Springer Nature och Taylor & Francis, i det att de båda har flera separata policyer, vilka stegvis skärper kraven på författarnas datahantering. Bland annat delning av data, data citering och DAS går från att uppmuntras i de första policyerna, till att bli ett krav. Olika policyer gäller även för olika tidskrifter, och således i förlängningen även olika forskningsdiscipliner (Springer Nature 2018b; Taylor & Francis 2018b). Intressant är att Taylor & Francis (2018b) har valt att uttrycka kraven på forskningsdatan genom FAIR-principerna, vilket tyder på en medvetenhet kring de teknologiska aspekterna av delad forskningsdata.

Speciellt för Springer Nature (2018b) är att de inkluderar instruktioner för peer review av publikationer med tillhörande data. Det är ett perspektiv som inte har dykt upp i vare sig de institutionella policyerna eller de från förlagen eller finansiärerna. Många tar på något sätt upp det ökade arbete och engagemang som krävs för att möjliggöra delad data och säkerställa en god datahantering, men få diskuterar den extra utmaningen som kommer med kvalitetsgranskningen av materialet. Ökad kvalitetsgranskning kan emellertid ses som en bra sak, eftersom ökad transparens och reproducerbarhet är viktiga argument för delad data, men då är det också bra om det utvecklas riktlinjer för hur dessa granskningar ska gå till.

Wiley (2018) använder sig också av en liknande stegvis policy, där delning av data uppmuntras, förväntas eller krävs. Dock är det något oklart vilken reell skillnad det blir mellan det första och det andra steget – om forskare uppmuntras eller förväntas dela sin data.

Elsevier (2018a) sticker ut från de andra förlagen i det att de knappt ställer upp några krav alls, utan lägger mer fokus på stöd till författare och på att främja öppen data. Detta kan tolkas olika – som en relativt fri policy som i stället fokuserar på att uppmuntra datadelning och god forskningsdatahantering, eller som ett exempel på ett svagt ställningstagande med oklara riktlinjer. Det har visats att forskare är obenägna att dela sin data om de inte är tvungna eller vinner något på det själva (Borgman 2012, Kiley 2013). Frågan är då om denna policy mest är till för att motivera författarna till att dela data. En annan förklaring kan ligga i de hänvisningar till de enskilda tidskrifterna som görs i policyn. Den övergripande datapolicyn kanske är mer generell, eftersom policyerna hos tidskrifterna är desto mer detaljerade.

Till sist kan nämnas att Springer Nature (2018a), Taylor & Francis (2018a) och Wiley (2018) tar upp att hänsyn tas till känslig data som av vissa anledningar inte kan delas, till exempel persondata, vilket går i linje med de allra flesta av de institutionella policyerna. Intressant är emellertid att inga andra undantag som är vanliga i de institutionella policyerna tas upp, som till exempel de kommersiella. Detta skulle kunna innebära konflikt i vissa fall. Dock är förlagens policyer olika strikta, så det kan hända att sådana data som har ett högt kommersiellt värde sällan publiceras i de tidskrifterna som har en strikt policy kring öppen data.

Bland finansiärerna har Wellcome Trust (2017) klart striktare krav på öppenhet, och en vidare definition av data och forskningsresultat än de andra finansiärerna som ingår i studien. Detta kan dels bero på att policyn har funnits längre än de andra, men även på att den verkar inom ett område som direkt påverkar allmänhetens väl och ve. De etiska och ideologiska argumenten blir än starkare när det kommer till medicinsk forskning. I kontrast till detta kan NWO (2018) nämnas, som både har en lösare definition av forskningsdata, och lättare krav gällande vilken data som ska sparas. NWO kräver endast att relevant data, som är lämplig för återanvändning, sparas, och beviljar finansiering för hantering av endast sådan relevant data. Här skulle en möjlig konflikt kunna uppstå, om institutionen eller tidskriften kräver en högre grad av öppenhet, och forskaren då inte kan söka finansiering för den extra datan. Intressant är även att NWO tar upp undantag som inte nämndes i förlagens policyer, till exempel kommersiella intressen.

Vetenskapsrådet (2015) baserar många av sina krav på lagstiftning, exempelvis lärosätenas arkivansvar och offentlighetsprincipen. Även undantagen grundas i lagstiftning, exempelvis personuppgiftslagen. Sådana riktlinjer fungerar egentligen mer som information, eftersom de forskare som dessa riktlinjer gäller även omfattas av den nämnda lagstiftningen. Riktlinjerna från Horizon 2020 är visserligen till viss del frivilliga, på grund av avskrivningsalternativet till ORD Pilot. Emellertid är ORD Pilot standard inom Horizon 2020 och forskarna förväntas göra sina data öppet tillgängliga (European Commission 2017).

Endast utifrån policyerna förefaller det egentligen inte finnas några stora problem att kombinera kraven och riktlinjerna från finansiärerna och förlagen. Det verkar gå bra att göra datan mer öppen än policyn kräver, så länge det inte bryter mot lagar eller etiska riktlinjer. Taylor & Francis (2018a) förstärker detta intryck när de hänvisar författare som funderar över kompatibiliteten mellan sina aktuella finansiärer och den tidskrift de ska publicera i, till SHERPA Juliet (2018). I de fall där tidskriften och finansiären inte stämmer överens uppmanas forskaren följa den policy som kräver störst öppenhet. Wiley (2018) tar det ett steg längre och erbjuder ett eget kompatibilitetsverktyg. Viktigt att komma ihåg är emellertid att vad som framkommer i policyerna inte nödvändigtvis speglar praktiken. Hur enkelt det är för forskare eller forskningsinstitutioner att tillgängliggöra data bestäms inte enbart av kompatibiliteten mellan policyerna, utan även av andra tekniska och sociala aspekter. Dessutom är de policyer som har studerats här översiktliga, och problem i kompatibiliteten kan mycket väl uppstå på individnivå. Till exempel har förlagen ofta olika bestämmelser för olika tidskrifter. Vissa kombinationer av forskningsdisciplin, finansiär och tidskrift kan visa sig inte vara kompatibla alls.

Som en sammanfattning kan sägas att de institutionella policyerna är anpassade efter önskemålen från externa aktörer, och utformade så att finansiärens krav står högst bland de riktlinjer som anges i policyerna. Det finns en sorts hierarkisk ordning mellan aktörerna, och relationerna dem emellan påverkar innehållen och utformningarna av datahanteringspolicyerna. Institutionella policyer – dess innehåll och utformning, påverkas uppenbarligen starkt av externa aktörer, och då särskilt av forskningsfinansiärer.

## 7 Diskussion

Syftet med studien var att undersöka vilka möjligheter och hinder det finns för öppen och delad data. Detta gjordes genom att jämföra ett antal officiella datahanteringspolicyer från europeiska universitet, för att se vilka aspekter av öppen data som tas upp och hur. Analysen av forskningsdatapolicyerna liknar på många sätt den av de nationella policyerna utförd av DCC och SPARC Europe (2018), vilken beskrevs i kapitel 2, men tar ner diskussionen på en institutionell nivå.

Nedan följer en kort sammanfattning av vad som har framkommit ur studien utifrån de forskningsfrågor som presenterades i kapitel 1.

*Vilka riktlinjer och krav gällande öppen data framställs av lärosäten, forskningsfinansiärer samt förlag, och hur skiljer sig dessa mellan de olika aktörerna?*

Det råder en allmänt positiv inställning till öppen data bland de universitet som ingick i studien. Deras policyer uttrycker en vilja att främja delning av forskningsdata, och en uppmaning till forskare att hantera sin data enligt god forskningssed. Hur tydligt detta uttrycks, exempelvis genom konkreta riktlinjer, varierar dock mellan olika universitet och kan ofta vara mycket vagt. All relevant data som underbygger vetenskapliga publikationer ska sparas i ett lämpligt arkiv och förses med tillräcklig metadata för att de ska kunna hittas och tolkas. Forskarna råds att tillgängliggöra sin data, men mycket fokus läggs även på undantagen från riktlinjerna, till exempel juridiska, etiska eller kommersiella skäl att inte göra datan öppen. Förlagen har en mer informativ inställning, och deras policyer klargör för vilka regler som gäller för olika tidskrifter. Ofta har förlagens policyer en stegvis karaktär som går från endast enkla riktlinjer kring datahantering till striktare krav på öppen data och korrekt dataåteranvändning. Finansiärerna har en liknande inställning till öppen data som universiteten, men mer konkreta riktlinjer och krav på hur och när data ska göras öppet tillgänglig, och vilken data det gäller.

*Utöver universitet, förlag och finansiärer – vilka andra aktörer inom forskarvärlden har betydelse för datadelning och -citering?*

Mycket ansvar för datahanteringen och tillgängliggörandet av data läggs på den enskilde forskaren och forskargruppen, trots att de själva vid första anblick inte har mycket att säga till om i frågan. Skiftningen till en mer öppen datahantering hänger i stor grad på forskarna, eftersom det är de som måste arbeta med att tillgängliggöra sin data. Ytterligare en aktör som har visat sig ha en stor roll, om än inte alltid uttalat i policyerna, är biblioteket. Mycket i de institutionella policyerna inriktas på stöd kring publicering och datahantering, vilket ofta faller på biblioteken att lösa. En tredje, mycket betydande aktör är staten, vars inflytande exempelvis visar sig i form av anpassningar till gällande lagar och regler samt bestämmelser kring fördelning av forskningsmedel. Detta kan ses på nationell nivå, men även på internationell nivå, till exempel inom EU. Till sist kan de aktörer nämnas som har att göra med marknaden för forskningsresultat. Till exempel tas kommersiella intressen upp genomgående i policyerna – mest i de institutionella, men även hos finansiärerna.

*Vilka motiveringar finns för dessa krav och riktlinjer, och vad säger det om de inblandade intressenterna?*

En av de vanligaste motiveringarna är allmänhetens intresse – att det som är offentligt finansierat ska komma allmänheten till nytta. Andra motiveringar relaterar till rådande lagar och regler. Universiteten hänvisar särskilt till krav från forskningsfinansiärer. Det ekonomiska perspektivet är också framträdande i policyerna – både gällande vem som ska betala för datahanteringen men också vilka ekonomiska intressen som kan tänkas finnas för datan som forskningsresultat.

## **7.1 Reflektion över analysen**

Ett av de viktigaste ämnena som tas upp är datapolicyernas syften, eftersom det går att koppla till alla de andra ämnena som inkluderas i analysen. En policy kan ändå sägas vara till för att uppnå något – för att få till någon form av förändring. Vid utformningen av policyn är det därför viktigt att det står klart vilka målen är, och hur de kan tänkas uppnås. Det är också viktigt att vara medveten om de effekter som policyn kan tänkas få – i detta fall effekterna på både de teknologiska och de sociala aspekterna av forskningsdatahanteringen. Ämnet passar väl in i överlappningen mellan den sociotekniska teorin och systemteorin, vilken riktar in sig på hur de olika komponenterna i ett system, och deras mål, påverkar varandra såväl som andra delar av systemet och systemet i stort (Golden 2013; Vancouver 2013).

Något som är talande för policyns syfte är vem den riktar sig till och vem som är avsändaren. I ljuset av diskussionen om mål och strävan efter förändring blir denna aspekt extra intressant. I denna studie har tre typer av avsändare inkluderats – nämligen universitet, förlag och finansiärer. Deras policyer riktar sig alla på något sätt till forskare, men med olika perspektiv. Förlagen har främst ett fokus på information kring den egna verksamheten och hur forskarna kan få tillgång till denna. I mycket ser det ut som ett kundperspektiv. I de institutionella policyerna märks ett tydligare forskarperspektiv och ett fokus på stöd i datahantering. Syftet blir att hjälpa fram både den enskilde forskaren och forskningen i stort, genom att förändra och förbättra den egna praktiken, och genom att hålla uppsikt över och anpassa sig efter externa aktörer. Universitetet, och särskilt universitetsbiblioteket, blir en knutpunkt i forskningsdatainfrastrukturen, vilket även återspeglas i datapolicyerna. Finansiärerna fokuserar mer på själva forskningen och dess öppenhet. I stället för att centreras kring verksamheten eller forskaren, utgår dessa mer ifrån allmänhetens perspektiv och intressen.

Ytterligare ett exempel på vad som kan bidra till syftet utan att det uttryckligen formuleras så i policyn är bestämmelserna kring revidering och utvärdering. Flera av policyerna har fastställt när de ska ses över och eventuellt ändras. Vanligast är årligen, men många av policyerna har ingen information om revidering alls. Att planera in en regelbunden uppdatering visar på en medvetenhet om forskningsdatans föränderlighet i forskarvärlden idag. Det är ett stort system med många aktörer, och det är bra att hålla sig uppdaterad. Ett större kontinuerligt engagemang från universitetens sida kan också ge större möjligheter för dem att påverka den egna rollen i forskningsdataekosystemet.



Många av punkterna i analysen anknyter även till frågan om tydlighet. Många policyer är tämligen oklara kring flera av riktlinjerna rörande datahantering. Som en första reflektion kan sägas att det verkar rätt meningslöst att fastställa en policy som ändå inte klargör vad som ska gälla. Om policyn endast beskriver en vision, men lämnar detaljerna otydliga, så kan tolkningen av den bli vad som helst. Det ger intrycket av att policyn endast har tagits fram eftersom det finns ett krav på att en policy ska finnas. En medvetet otydlig formulering kan också vara ett sätt att enkelt sätta punkt för en komplicerad fråga. Emellertid skulle det även kunna finnas en poäng i att uttrycka sig luddigt, genom något sorts outtalat syfte. Frågan är hur institutionen tror sig kunna uppnå sin målbild. Är det bäst att ställa krav och sätta regler, eller att erbjuda stöd och uppmuntra till delning, och i stället informera om de krav som ställs från förlag och finansärer? Troligtvis beror behovet av krav och regler på hur långt de som berörs av policyn har kommit i arbetet med öppen data, och vilken sorts data som det rör sig om. Om institutionen som antar policyn till exempel endast bedriver forskning som hanterar stora mängder känsliga persondata, så finns det inte lika stor anledning att kräva datadelning av forskarna. En inställning att data bör delas när det är möjligt är kanske mer passande, och fullt tillräcklig, i sådana fall.

Universitetens syften och mål med datahanteringspolicyerna hör även ihop med hur den grundläggande synen på data tar sig uttryck. Forskningsdata kan beskrivas som ett problem som behöver lösas, outtalat eller uttalat, eller mer som en tillgång eller möjlighet till utveckling. Överlag verkar det senare gälla i de institutionella policyerna. I definitionerna dominerar de relationella delarna, det vill säga datans förmedlande egenskaper. Fokus ligger på data som är användbara och relevanta, och som underbygger publicerade resultat. Detta syns även om man ser till de olika kraven på öppen data som presenteras, där fokus läggs på nyttan med datan. Emellertid kan även kopplingar till data som ett problem påträffas. Den ekonomiska biten tas upp av majoriteten. Vem som ska betala för datahanteringen är utan tvivel en viktig fråga. En annan relevant fråga är den om metadata. Det understryks att god datahantering kräver tillräcklig och högkvalitativ metadata, vilket i sin tur kräver en hel del arbete och tid. Fler kopplingar till problem kan återfinnas i motiveringarna för öppen data. Till exempel anges krav från finansärer som universiteten och forskarna har att anpassa sig efter. De hänvisar även i många fall till nationella riktlinjer för öppen data. Även om universitetet delar uppfattningen, så framstår det ändå som ett problem som måste hanteras. De teknologiska aspekterna tas endast upp av ett fåtal, till exempel Newcastle University (2017) som anger hur mycket lagringsutrymme som erbjuds forskningsprojekten. De teknologiska lösningarna kan onekligen innebära problem med datahanteringen, men i de flesta fall verkar det glömmas bort i policyerna. Det kan bero på att det ses som underförstått, eller som att det inte passar in i en policy, men det skulle också kunna bero på okunskap om de teknologiska systemen och hur de bör utvecklas.

Forskningsdata som problem eller möjlighet väcker en intressant diskussion om forskningsstöd och forskningsstyrning. Att lösa datahanteringsproblem lokalt och hjälpa till med tillgängliggörandet är tydliga exempel på forskningsstöd. Emellertid kan problemlösning också tolkas som styrning, till exempel om man ser till vad som ses som problem och hur dessa avhjälpas. Den anpassning som sker efter externa intressen och

riktlinjer är också exempel på forskningsstyrning. Frågan blir då vem det är som styr. Vilka relationerna är mellan aktörerna och vad de har för intressen i problemet i fråga, påverkar datahanteringen och policyernas utformning. Universitetet skiftar här konstant roll, eller snarare så innehar det en dubbel roll – en stödjande och en styrande.

Till sist kan nämnas något om strävan efter förändring och trögheten i forskningsdatasystemet. Ovan nämndes att policyer i princip är till för att uppnå något sorts mål. Vissa policyer visar tydligt att en förändring är önskvärd – oftast att forskning och forskningsdata ska bli mer öppen – men att det finns ett antal hinder som bromsar denna utveckling. Denna tröghet kan härledas till lagar och kommersiella intressen, men även till forskarna själva. Det handlar om ett stort system där inblandade aktörer är vana vid att publicering och datahantering går till på ett visst sätt, och det går därför långsamt att förändra. I mycket liknar det de självreglerande system som beskrivs av Vancouver (2013). Forskarna måste själva lägga ner tid och arbete på datahanteringen, vilket kräver någon sorts incitament eller uppmuntran. Det är mer sannolikt att data delas om processen för tillgängliggörandet är problemfri och automatiserad (Borgman 2012), men forskarna kan inte uteslutas helt. Till exempel är de med i urvalsprocessen. Det är ingen mening med att slentrianmässigt spara all data, utan en bedömning måste göras av någon som förstår datan – vad som är värdefullt och i vilken form det ska sparas. En liknelse kan göras med de institutionella policyerna och deras syften, som ibland kan vara mycket vaga. Det är nämligen heller ingen mening med att skriva en policy endast för att det måste finnas en policy. Universitetet skulle kunna gynnas av att få sina syften och mål med datahanteringen klara för sig själva, och tydliggöra detta, för vad är det egentligen som menas med att data ska tillgängliggöras enligt god sed, eller inom rimlig tid?

## **7.2 Reflektion över studien**

Den genomförda studien utgörs av en dokumentanalys, som grundar sig i koncept från socioteknisk teori och systemteori. Den valda teorin har framförallt fungerat som riktlinjer för vilka frågor som skulle stå i fokus i analysen, och har då varit fullt tillräckliga. Det som teorin i huvudsak har bidragit med till studien är en medvetenhet kring samspelet mellan de tekniska och sociala komponenterna i forskningsdataekosystemet. De problem och fördelar med öppen forskningsdata som framkommer i policyerna är av såväl teknologisk karaktär, till exempel rörande lagringsutrymme och infrastruktur, som av social, rörande exempelvis akademiskt genomslag, transparens inom forskning samt incitament till datadelning. En förändring är till synes på gång, men ett flertal svårigheter skapar hinder. Detta kan med fördel liknas vid systemteorins kontrollsystem och tröghet vid förändringar. En annan viktig aspekt från teorin – den sociotekniska delen – är vikten av insyn i verksamheten och systemet för dem som arbetar med och inom det. För att ha möjlighet att påverka och utveckla måste universitetet och forskarna själva ha förståelse för systemet. Ytterligare en parallell till det sociotekniska perspektivet kan dras från konceptet Open Science, som på sätt och vis förstärker skiftet från den traditionella forskningspubliceringens löpande band-karaktär till ett mer individanpassat och transparent publiceringssystem.

I analysen behandlades synen på forskningsdata, och huruvida den ses som problem eller

möjlighet, vilket anknyter till systemteorin och dess beskrivning av systemens problemlösningsfunktion. Datahanteringen och -delningen som system kan tänkas vara till för att lösa problem med tillgänglighet, transparens och möjligheter att återskapa forskning. Även synen på informationsutbyte som en urvalsprocess har betydelse för policyernas syften, genom frågor om vilka data som delas och förmedlas, och vad detta kommunicerar till omgivningen. Tätt kopplat till kommunikation och informationsutbyte ligger frågan om bibliotekens roll. Inom både den sociotekniska teorin och systemteorin betonas systemets interaktion med omgivningen och hur detta påverkar systemet i fråga. Biblioteket kan här sägas inta en gränsöverskridande roll med syfte att balansera de olika delar av forskningsdataekosystemet med omgivningen. Detta perspektiv kan ge en bra startpunkt både för analys och utformning av institutionella datahanteringspolicyer, vilket biblioteken ofta är inblandade i.

Materialet som studien baseras på är visserligen öppet och lättillgängligt, men i och med att det till stor del är internetbaserat, och dessutom i många fall kontinuerligt uppdaterat kan det också ses som förgängligt material. Det kan i sådana fall vara svårt att återkomma till det fullständiga ursprungsmaterialet i efterhand. Dock sker uppdateringarna som mest årligen, och många av policyerna är nya och har aldrig reviderats. Vissa policyer har även en revideringshistorik, vilken visar att inte mycket har ändrats från ursprungsdokumentet. I en övergripande studie av mönster och trender bland datapolicyer bör förgängligheten inte skapa något problem. För denna studie finns emellertid en del av rådatan, det vill säga policyerna, sparade i den version som ligger till grund för undersökningen. Alla institutionella policyer, förutom de från University of Warwick, University of York och Stockholms universitet, är sparade hos författaren och finns tillgängliga vid förfrågan. Detta gäller även rapporterna från Vetenskapsrådet och Horizon 2020. De andra policyerna, inklusive de från förlagen, är helt webbaserade. Trots att en stor del av materialet sparas, finns det få tekniska möjligheter att tillgängliggöra det samlat i anslutning till uppsatsen. För masteruppsatser finns det i nuläget heller inga krav på att bakomliggande data ska tillgängliggöras.

Att materialet till viss del är lättillgängligt och öppet tillför en transparens till studien, och det är lätt att följa vad som har gjorts och hur. Den valda teorin och metoden är också relativt enkla och lättillgängliga, vilket går väl ihop med diskussionen om tydlighet, som återkommer i flera av studiens delar. Dokumentanalysen som metod fungerar bra för en sådan här studie, eftersom det är stort fokus på primärmaterialet, som är tydligt definierat. Dokumentanalysen används emellertid ofta i kombination med någon annan metod och material för att kunna bekräfta resultaten genom triangulering (Gibson & Brown 2011). I denna studie användes visserligen dokumentanalysen i kombination med en komparativ metod, vilket gav ett lite bredare perspektiv till analysen, men en utveckling av metoden kan ändå vara värdefull. En idé för att bygga en mer fullständig bild av forskningsdatans förutsättningar, och för att få en uppfattning om hur väl det skrivna i institutionella policyer stämmer med den verkliga praktiken, vore att inkludera någon form av intervju- eller enkätundersökning. På det sättet skulle även ett forskarperspektiv kunna vägas in i studien.

### 7.3 Framtida forskning

Denna studie, eller liknande, kan användas som bakgrund när en institutionell datahanteringspolicy ska utvecklas. Vid ett sådant arbete är det bra att utvärdera hur policyerna ser ut vid andra universitet. Syftet är särskilt viktigt att reflektera över, eftersom själva anledningen till att skriva en policy är att något ska uppnås och att målen behöver struktureras. Det kan också vara värdefullt att se vilka ämnen som behandlas i andra institutionella policyer och hur det bidrar till helheten, eftersom det kan användas som underlag vid besluten gällande vilka ämnen som ska tas upp i den egna policyn.

Det har visat sig finnas en del forskningsluckor inom ämnet öppen data och forskningsdatahantering, och även en del intressanta utvecklingsmöjligheter. Till exempel behövs det utföras mer forskning kring återanvändning av data och datacitering – hur användningen ser ut, vilka standarder som förespråkas och deras för- och nackdelar, samt dess verkningar inom forskarvärlden. Den allmänna uppfattningen, i alla fall inom bibliotekssfären, verkar vara att öppen data är något bra, men det problematiseras inte i någon stor utsträckning. Många studier fokuserar på akademiskt genomslag och citeringsanalyser. Intressant hade varit att se mer kvalitativa studier kring hur forskare ställer sig till att använda sig av andras insamlade data, samt hur forskningen påverkas av det.

Ett annat intressant utvecklingsområde rör forskningsdatans kontextberoende och skillnaden mellan olika forskningsdiscipliner. En fortsättning på studien vore att göra disciplinspecifika studier av kraven på öppen data.

Det finns mycket att utveckla och studera inom ämnet forskningsdata, av vilket endast en liten del har tagits upp i denna uppsats. Datahantering och öppen tillgång till forskningsdata är ämnen som blir alltmer aktuella – särskilt inom biblioteks- och informationsvetenskapen. Detta beror på teknologiska orsaker, så som den ökande mängden forskningsdata, såväl som sociala eller politiska, som de ökande kraven på statligt finansierade forskningsresultat ska göras öppet tillgängliga. Detta märks även av i Sverige, där arbetet med att utveckla institutionella datahanteringspolicyer pågår vid flera forskningsinstitutioner, och fler samarbeten gällande datahantering formas institutioner emellan.

Mitt i allt detta befinner sig forsknings- och högskolebiblioteket. Trots sin funktion som naturlig samlings- och knypunkt för forskarvärlden, är bibliotekets roll tämligen oklar, vilket till exempel visade sig i Briney, Goban och Zilinskis (2015) studie. Från att ha varit främst bevarande och tillgängliggörande har bibliotekets roll i och med digitaliseringen utvecklats och utvidgats till nya områden. Forskningsdatahantering är inget undantag. Det är möjligt att universitetet såväl som biblioteken skulle vinna på att tydliggöra bibliotekens roll i frågan, men även på att undersöka och ta till vara på den stora potential som finns hos biblioteken gällande forskningsdatahantering.

## Referenser

Appelbaum, S. H. (1997). Socio-technical systems theory: an intervention strategy for organizational development. *Management Decision* 35(6) ss. 452–463. doi:10.1108/00251749710173823.

Ayris, P. (2013). *Research Data Policy*. Version 4.0. University College London. URL: <http://www.ucl.ac.uk/library/research-support/research-data/policies> [2018-03-16].

Baker, H. & Clay, D. (2016). *Research Data Management Policy*. Version 1.0. University of Salford. URL: <http://www.salford.ac.uk/about-us/corporate-information/governance/policies-and-procedures/library-and-information-resources> [2018-03-16].

Bartling, S. & Friesike, S. (2014). Towards Another Scientific Revolution. I Bartling, S. & Friesike, S. (red.) *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*. Cham: Springer International Publishing. Kap. 1 ss. 3–15. doi:10.1007/978-3-319-00026-8\_1.

Borgman, C. L. (2012). The Conundrum of Sharing Research Data. *Journal of the American Society for Information Science and Technology* 63(6) ss. 1059–1078. doi:10.1002/asi.22634.

– (2015). *Big data, little data, no data: scholarship in the networked world*. Cambridge, Massachusetts: The MIT Press.

Bowen, G. A. (2009). Document Analysis as a Qualitative Research Method. *Qualitative Research Journal* 9(2) ss. 27–40. doi:10.3316/QRJ0902027.

Brase, J., Socha, Y., Callaghan, S., Borgman, C. L., Uhler, P. F. & Carroll, B. (2014). Data Citation: Principles and Practice. I Ray J.R. (red.) *Research data management: practical strategies for information professionals*. West Lafayette, Indiana: Purdue University Press. Kap. 8 ss. 167–186.

Briggs, M. (2016). *Research Data Management Policy*. Version 0.2. University of Wales Trinity Saint David. URL: <http://www.uwtsd.ac.uk/library/research-data-management/about-rdm/> [2018-03-16].

Briney, K., Goben, A. & Zilinski, L. (2015). Do You Have an Institutional Data Policy? A Review of the Current Landscape of Library Data Services and Institutional Data Policies. *Journal of Librarianship and Scholarly Communication* 3(2), eP1232. doi:10.7710/2162-3309.1232.

Brown, M. L. (2013). The role of the research library. I Shorley, D. & Jubb, M. (red.) *The Future of Scholarly Communication*. London: Facet Publishing. Kap. 12 ss. 157–168.

Buckinghamshire New University (2016). *Research Data Management Policy*. Version 1.0. Academic Quality Directorate. URL: <https://bucks.ac.uk/about-us/governance-and-policies/policies> [2018-03-16].

Buneman, P., Davidson S. & Frew, J. (2016). Why Data Citation Is a Computational Problem. *Communications of the ACM* 59(9) ss. 50–57. doi:10.1145/2893181.

Chiang T. (2016). *Stories of Your Life and Others*. New York: Vintage Books.

CODATA (2018). *CODATA's Mission*. International Council for Science (ISCU). URL: <http://www.codata.org/about-codata/our-mission> [2018-03-12].

DataCite (2018a). *Members*. URL: <https://www.datacite.org/members.html> [2018-03-12].

– (2018b). *Our Mission*. URL: <https://www.datacite.org/mission.html> [2018-03-12].

DCC & SPARC Europe (2018). *An Analysis of Open Data and Open Science Policies in Europe*. Version 2.1. URL: <https://sparceurope.org/new-sparc-europe-report-analyses-open-data-open-science-policies-europe/> [2018-03-30].

Digital Curation Centre (2018). *About the DCC*. URL: <http://www.dcc.ac.uk/about-us> [2018-03-12].

Drachen, T., Ellegaard, O., Larsen, A. & Dorch, S. (2016). *Sharing data increases citations*. *Liber Quarterly* 26(2) ss. 67–82. doi:10.18352/lq.10149.

Edge Hill University (2017). *Research data management policy*. URL: <https://www.edgehill.ac.uk/documents/department/research-office/> [2018-03-16].

Ekström, B. (2017). Vart är open access-rörelsen på väg? Fyra tendenser i svenska & brittiska open access-nätverk. *Tidskrift för ABM* 2(2) ss. 5–35.

Elsevier (2018a). *Research Data*. URL: <https://www.elsevier.com/about/our-business/policies/research-data> [2018-04-11].

- Elsevier (2018b). *Research Data FAQs*. URL: <https://www.elsevier.com/about/our-business/policies/research-data-faqs> [2018-04-11].
- European Commission (2017). *Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020*. Version 3.2. Horizon 2020. URL: [http://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf) [2018-03-14].
- (2018). *What is Horizon 2020?* Horizon 2020. URL: <https://ec.europa.eu/programmes/horizon2020/en/what-horizon-2020> [2018-03-14].
- Evans, J. (2015). *Research Data Management Policy*. Version 3.0. University of Central Lancashire. URL: <https://www.uclan.ac.uk/research/environment/strategy.php> [2018-03-16].
- Fairclough, A. (2017). *Policy on Research Data Management*. Version 04. RDMPLN-POL\_006. University of Liverpool. URL: <https://www.liverpool.ac.uk/library/research-data-management/> [2018-03-16].
- Fecher, B. & Friesike, S. (2014). Open Science: One Term, Five Schools of Thought. I Bartling, S. & Friesike, S. (red.) *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*. Cham: Springer International Publishing. Kap. 2 ss. 17–47. doi:10.1007/978-3-319-00026-8\_2.
- Gibson, W. J. & Brown, A. (2011). Using Documents in Research. *Working with Qualitative Data*. London: SAGE Publications. Kap. 5 ss. 65–83.
- Golden, T. D. (2013). Sociotechnical Theory. I Kessler E. H. (red.) *Encyclopedia of Management Theory*. Thousand Oaks: SAGE Publications, Ltd. ss. 752–755.
- Heylighen, F. (1998). *Basic Concepts of the Systems Approach*. Principia Cybernetica Web. URL: <http://pespmc1.vub.ac.be/SYSAPPR.html> [2018-06-18].
- Heylighen, F. & Joslyn, C. (1992). *What is Systems Theory?*. Principia Cybernetica Web. URL: <http://pespmc1.vub.ac.be/SYSTHEOR.html> [2018-06-18].
- Higman, R. & Pinfield, S. (2015). Research data management and openness: The role of data sharing in developing institutional policies and practices. *Program: electronic library and information systems* 49(4) ss. 364–381. doi:10.1108/PROG-01-2015-0005.
- Horton, L. & DCC (2016). *Overview of UK Institution RDM Policies*. Digital Curation Centre. URL: <http://www.dcc.ac.uk/resources/policy-and-legal/institutional-data-policies> [2018-03-30].

Jubb, M. (2013). Introduction: Scholarly communications – disruptions in a complex ecology. I Shorley, D. & Jubb, M. *The Future of Scholarly Communication*. London: Facet Publishing ss. XIII–XXXVI.

Kiley, R. (2013). The view of the research funder. I Shorley, D. & Jubb, M. (red.) *The Future of Scholarly Communication*. London: Facet Publishing. Kap. 10 ss. 131–144.

Knight, G. (2014). *LSHTM Research Data Management Policy*. London School of Hygiene & Tropical Medicine, URL:  
<https://www.lshtm.ac.uk/research/research-publications-research-data> [2018-03-16].

Lancaster University (2015). *Research Data Policy*. Version 2. SEC/2015/2/0416. URL:  
<http://www.lancaster.ac.uk/library/rdm/what-is-rdm/> [2018-03-16].

Leonelli, S. (2015). What Counts as Scientific Data? A Relational Framework. *Philosophy of Science* 82(5) ss. 810–821. doi:10.1086/684083.

Mauthner, N. S. & Parry, O. (2013). Open Access Digital Data Sharing: Principles, Policies and Practices. *Social Epistemology* 27(1) ss. 47–67.  
doi:10.1080/02691728.2012.760663.

Morrison, H. (2009). *Scholarly communication for librarians*. Oxford: Chandos, ss. 103–134.

Newcastle University (2017). *Research Data Management Policy Principles & Code of Good Practice*. University of Newcastle Upon Tyne. URL:  
<https://research.ncl.ac.uk/rdm/beforeaproject/expectations/> [2018-03-16].

NWO (2018). *Data management protocol*. URL:  
<https://www.nwo.nl/en/policies/open+science/data+management> [2018-04-13].

OpenAIRE (2017). *Project factsheets: The OpenAIRE2020 Project*. URL:  
<https://www.openaire.eu/project-factsheets> [2018-03-12].

Proposition 2016/17:50. *Kunskap i samverkan – för samhällets utmaningar och stärkt konkurrenskraft*. Stockholm: Utbildningsdepartementet. URL:  
<https://www.regeringen.se/4adad0/contentassets/72faaf7629a845af9b30fde1ef6b5067/kunskap-i-samverkan-for-samhallets-utmaningar-och-starkt-konkurrenskraft-prop.-20161750.pdf> [2018-05-21].

Research Data Alliance (2017). *Who is RDA?* URL:  
<https://www.rd-alliance.org/about-rda/who-rda.html> [2018-03-12].



- Research Data Alliance (2018). *RDA in a nutshell*. URL: <https://www.slideshare.net/ResearchDataAlliance> [2018-03-12].
- Rice, R. & Southall, J. (2016). *The data librarian's handbook*. London: Facet Publishing.
- ROARMAP (2018). *Registry of Open Access Repository Mandates and Policies*. URL: <http://roarmap.eprints.org/> [2018-03-12].
- Rzepa, H. S. (2013). Changing ways of sharing research in chemistry. I Shorley, D. & Jubb, M. (red.) *The Future of Scholarly Communication*. London: Facet Publishing. Kap. 1 ss. 3–23.
- Sawyer, S. & Jarrahi, M. H. (2013). *Sociotechnical approaches to the study of Information Systems*. Syracuse University. URL: <http://sawyer.syr.edu/publications/2013/sociotechnical%20chapter.pdf> [2018-06-15].
- SHERPA (2018). University of Nottingham. URL: <http://www.sherpa.ac.uk/> [2018-03-12].
- SHERPA Juliet (2018). *Research Funders' Open Access Policies*. URL: <http://v2.sherpa.ac.uk/juliet/> [2018-03-12].
- SHERPA/FACT (2018). *Funders & Authors Compliance Tool*. URL: <http://www.sherpa.ac.uk/fact/> [2018-03-12].
- SHERPA/RoMEO (2018). *Publisher copyright policies & self-archiving*. URL: <http://www.sherpa.ac.uk/romeo/index.php> [2018-03-12].
- Sidler, M. (2014). Open Science and the Three Cultures: Expanding Open Science to all Domains of Knowledge Creation. I Bartling, S. & Friesike, S. (red.) *Opening Science: The Evolving Guide on How the Internet is Changing Research, Collaboration and Scholarly Publishing*. Cham: Springer International Publishing. Kap. 5 ss. 81–85. doi:10.1007/978-3-319-00026-8\_5.
- Springer Nature (2018a). *Research Data Policies FAQs*. URL: <https://www.springernature.com/gp/authors/research-data-policy/faqs/12327154> [2018-04-11].
- (2018b). *Research Data Policy Types*. URL: <https://www.springernature.com/gp/authors/research-data-policy/data-policy-types/12327096> [2018-04-11].

Stephenson, T. & Everitt, M. (2016). *Management of Research Data Policy*. Version 3. CU-RIO-POL-4.0. Cranfield University, Research and Innovation Office. URL: <https://www.cranfield.ac.uk/about/governance-and-policies/policies-and-Regulations> [2018-03-16].

Stichweh, R. (u.å.). *Systems Theory*. University of Bonn. URL: [https://www.fiw.uni-bonn.de/demokratieforschung/personen/stichweh/pdfs/80\\_stw\\_systems-theory-international-encyclopedia-of-political-science\\_2.pdf](https://www.fiw.uni-bonn.de/demokratieforschung/personen/stichweh/pdfs/80_stw_systems-theory-international-encyclopedia-of-political-science_2.pdf) [2018-06-18].

Stockholms universitet (2018). *Forskningsdatapolicy för Stockholms universitet*. Dnr SUFV 5.1.1-1780-17. URL: <https://www.su.se/biblioteket/publicera/forskningsdata/riktlinjer/forskningsdatapolicy-f%C3%B6r-stockholms-universitet-1.376329> [2018-04-10].

Suber, P. (2012). *Open Access*. Cambridge, Mass.: MIT Press. URL: <https://mitpress.mit.edu/sites/default/files/titles/content/openaccess/cover.html> [2018-03-07].

Sveinsdottir, T., Wessels, B., Smallwood, R., Linde, P., Kala, V., Tsoukala, V. & Sondervan, J. (2013). *Stakeholder Values and Ecosystems*. Policy RECommendations for Open access to research Data in Europe (RECODE). Seventh Framework Programme for Science in Society.

Svensk nationell datatjänst (2018). *Om oss*. URL: <https://snd.gu.se/sv/om-oss> [2018-03-12].

Taylor & Francis (2018a). *Data sharing FAQs*. URL: <https://authorservices.taylorandfrancis.com/data-sharing-faqs/> [2018-04-11].

– (2018b). *Data sharing policies*. URL: <http://2qkk0e1599xt254aernh2gta-wpengine.netdna-ssl.com/wp-content/uploads/2018/01/Data-sharing-policies.pdf> [2018-04-11].

– (2018c). *Understanding our data sharing policies*. URL: <https://authorservices.taylorandfrancis.com/understanding-our-data-sharing-policies/> [2018-04-11].

Trist, E. L. & Bamforth, W. (1951). Some Social and Psychological Consequences of the Longwall Method of Coal-Getting. *Human Relations* 4(1) ss. 3–38. doi:10.1177/001872675100400101.

University of Bristol (2013). *Research data management and open data policy*. URL: <http://www.bristol.ac.uk/research/environment/governance/research-data-policy/> [2018-03-16].

University of Manchester (2015). *Research Data Management Policy*. Version 1.1. URL: <http://www.library.manchester.ac.uk/using-the-library/staff/research/services/research-data-management/policies/> [2018-03-16].

University of Warwick (2011). *Research Data Management Policy*. URL: [https://warwick.ac.uk/services/ris/research\\_integrity/code\\_of\\_practice\\_and\\_policies/research\\_code\\_of\\_practice/datacollection\\_retention/research\\_data\\_mgt\\_policy](https://warwick.ac.uk/services/ris/research_integrity/code_of_practice_and_policies/research_code_of_practice/datacollection_retention/research_data_mgt_policy) [2018-04-09].

University of York (2017). *Research Data Management Policy*. URL: <https://www.york.ac.uk/about/departments/support-and-45admin/information-services/information-policy/index/research-data-management-policy/#tab-1> [2018-04-04].

Utrecht University (2016). *University policy framework for research data*. URL: <https://www.uu.nl/en/research/research-data-management/guides/policies-codes-of-conduct-and-laws> [2018-04-10].

Vancouver, J. B. (2013). Systems Theory of Organizations. I Kessler, E. H. *Encyclopedia of Management Theory*. Thousand Oaks: SAGE Publications, Ltd. ss. 816–820.

Vetenskapsrådet (2015). *Förslag till Nationella riktlinjer för öppen tillgång till vetenskaplig information*. Vetenskapsrådets rapporter, 1.1.2-2015-8. URL: [https://publikationer.vr.se/produkt/forslag-till-nationella-riktlinjer-for-oppentillgang-till-vetenskaplig-information/?\\_ga=2.133174094.1190685332.1522740873-1732467253.1513865014](https://publikationer.vr.se/produkt/forslag-till-nationella-riktlinjer-for-oppentillgang-till-vetenskaplig-information/?_ga=2.133174094.1190685332.1522740873-1732467253.1513865014) [2018-04-15].

Wellcome Trust (2017). *Policy on data, software and materials management and sharing*. URL: <https://wellcome.ac.uk/funding/managing-grant/policy-data-software-materials-management-and-sharing> [2018-04-15].

– (2018). *Developing an outputs management plan*. URL: <https://wellcome.ac.uk/funding/managing-grant/developing-outputs-management-plan> [2018-04-16].

Wessels, B. et al. (2014). Issues in the development of open access to research data. *Prometheus* 32(1) ss. 49–66. doi:10.1080/08109028.2014.956505.

Whitley, R. (2011). Changing Governance and Authority Relations in the Public Sciences. *Minerva: A Review of Science, Learning and Policy* 49(4) ss. 359–385. doi:10.1007/s11024-011-9182-2.

Wiley (2018). *Sharing and Citing your Research Data*. URL:  
[https://authorservices.wiley.com/author-resources/Journal-Authors/  
licensing-open-access/open-access/data-sharing.html](https://authorservices.wiley.com/author-resources/Journal-Authors/licensing-open-access/open-access/data-sharing.html) [2018-04-11].

Wilkinson, M. D. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3 art.nr 160018.  
doi:10.1038/sdata.2016.18.

Willinsky, J. (2006). *The Access Principle: The Case for Open Access to Research and Scholarship*. Cambridge, Mass.: MIT Press.

Young, M., Sørensen, M. P., Carter Bloch, C. & Degn, L. (2017). Systemic rejection: political pressures seen from the science system. *Higher Education: The International Journal of Higher Education Research* 74(3) ss. 491–505.  
doi:10.1007/s10734-016-0059-z.