

Stine Granviken

---

**Institusjonelt arkiv - ny kanal for vitenskaplig  
og faglig produksjon**

Masteroppgave  
Avdeling for journalistikk, bibliotek og informasjonsfag

## **Sammendrag**

Oppgaven er en studie av publisering og tilgjengeliggjøring av vitenskaplig og faglig materiale i den norske universitets- og høgskolesektoren. Gjennom en kvantitativ undersøkelse er innføring og bruk av institusjonelt arkiv (IA) som ny formidlingskanal ved to universitet i Norge og to i Sverige analysert og diskutert. Hovedformålet har vært å undersøke om innføringen av IA har hatt en positiv effekt på gjenfinning og bruk av institusjonens forskningsproduksjon, dvs. om IA gir en verdiøkning.

I gjennomføring av undersøkelsen er en ny metode utviklet og testet. For å finne forskernes ”research impact” eller anerkjennelse setter den ”nye” metoden forskerne og deres produksjon i fokus i motsetning til å ta utgangspunkt i vitenskapelige tidsskrifter og artiklene i disse, som er en mer vanlig måte å undersøke anerkjennelse på. Den nye metoden sammenligner forfattere, faggrupper og institusjoner med seg selv, dvs. at siteringer for artikler som er parallellpublisert i IA og artikler som ikke er parallellpublisert fra samme forfatter sammenlignes. Dette kan skaleres slik at man kan sammenligne siteringer fra flere forfattere / faggrupper innen samme institusjon, eller fra flere forfattere / faggrupper fra flere institusjoner. Som basis brukes en tradisjonell bibliometrisk metode med statistiske analyser av faglige siteringer og referanser. Siteringer for et antall vitenskapelige artikler for et utvalg av forskere telles, og siteringshastighet analyseres og sammenlignes for artikler som er parallellpublisert i IA med artikler som kun er publisert i tradisjonelle kanaler. Antall siteringer aggregeres opp til det nivået som er hensiktsmessig for den undersøkelsen som skal gjennomføres. Her i denne oppgaven er det aggregert og foretatt sammenligninger på nivå for emnekategori, institusjon og land.

Tilgjengelig datagrunnlag for Norge er pr dato begrenset, både fordi IA som formidlingskanal er ny, og fordi rutiner knyttet til parallellpublisering fortsatt er under utarbeidelse og i innkjøringsfasen hos de fleste av institusjonene som har innført IA. Resultatene i denne undersøkelsen er derfor ikke eksakte fakta, men gir gode indikasjoner på utviklingstendenser for antall siteringer for artikler parallellpublisert sammenlignet med artikler som ikke er det.

Basert på gjennomføring av analyser av antall siteringer fra ulike vinkler og ulike tidsintervaller viser resultatene en utviklingstendens i retning av en større økning i antall siteringer for artikler som er parallellpublisert i IA sammenlignet med artikler som kun er publisert i tradisjonelle kanaler. Resultatene viser også at det er en viss sannsynlighet for raskere sitering av artikler som er parallellpublisert i IA enn artikler som ikke er parallellpublisert.

Trenden etter få år med institusjonelle arkiv i Norge tyder på at parallellpublisering kan gi bedre synlighet, flere siteringer, og dermed økt anerkjennelse.

**Kunnskap kan være to ting; Det vi vet, og det at vi vet hvor vi kan lære det vi ikke vet.**

Samuel Johnson, 1709-1784.

Masteroppgave ved Høgskolen i Oslo, Avdeling for journalistikk, bibliotek- og  
informasjonsfag  
Oslo 2009

# Innhold

<b>1. INNLEDNING</b> .....	<b>4</b>
1.1 OPEN ACCESS .....	5
1.2 INTERNASJONAL OG NASJONAL STATUS .....	7
1.3 INSTITUSJONELLE ARKIV - FORMÅL .....	11
<b>2. LITTERATUR OG TIDLIGERE FORSKNING</b> .....	<b>13</b>
2.1 INSTITUSJONELLE ARKIV I UTLANDET .....	13
2.2 INSTITUSJONELLE ARKIV I NORGE .....	14
2.3 INSTITUSJONELLE ARKIV OG ”IMPACT” .....	15
2.4 INSTITUSJONELLE ARKIV I ET STRATEGISK OG FORVALTNINGSMESSIG PERSPEKTIV .....	17
<b>3. FORSKNINGSSPØRSMÅL</b> .....	<b>20</b>
<b>4. PRAKTISK TILNÆRMING - METODE</b> .....	<b>21</b>
4.1 PROSJEKTETS MATERIALE .....	23
4.2 STATISTISK REPRESENTATIVITET .....	33
<b>5. PRESENTASJON AV RESULTATER - ANALYSE OG DISKUSJON</b> .....	<b>35</b>
5.1 ENDRINGER I ANTALL SITERINGER .....	36
5.1.1 Endringer i antall siteringer før og etter tidspunktet for innføring av IA .....	37
5.1.2 Endringer i antall siteringer for artikler parallellpublisert i IA eller ikke .....	42
5.2 SITERINGSHASTIGHET .....	46
5.2.1 Endringer i siteringshastighet for artikler parallellpublisert i IA eller ikke .....	46
5.2.2 ”Ny giv” – oppsving i antall siteringer for ”gamle” artikler .....	50
5.2.3 Årsfordeling av siteringer i perioden 1998-2008 .....	55
5.3 PROBLEMSTILLINGER OG PÅVIRKNINGSFAKTORER .....	61
5.4 ERFARINGER MED BRUK AV ”NY” METODE .....	65
<b>6. OPPSUMMERING OG KONKLUSJON</b> .....	<b>68</b>
<b>7. LITTERATURLISTE</b> .....	<b>74</b>
<b>8. VEDLEGG</b> .....	<b>78</b>
8.1 ORDNET LISTE MED OVERSIKT OVER FORFATTERE OG DERES ARTIKKELPRODUKSJON I NORA/DIVA ....	78
8.2 FORFATTERUTVALGET - UTGANGSPUNKT FOR SØK I ISI .....	79
8.3 ARTIKKELUTVALGET - UTTREKK FRA ISI .....	80
8.4 SAMMENSTILLING AV UTTREKK FRA NORA/DIVA OG ISI .....	81

# 1. Innledning

I løpet av det siste tiåret har det vært en rivende utvikling i bruk av digitale medier og Internett. Alle sektorer og bransjer har tatt dette i bruk på ulike måter, og biblioteksektoren er intet unntak. Bibliotekene står overfor store utfordringer, men også muligheter når det gjelder å ta i bruk digitale verktøy som supplerende og fremtidsrettede tjenester. Tradisjonelt har bibliotekets oppgave vært å oppbevare bøker og manuskripter og å ordne samlingene slik at de er tilgjengelige for brukerne. I dag består bibliotekenes samlinger i økende grad av digitale dokumenter, og oppgaven med samlingsoppbygging og kunnskapsorganisering er endret.

Snakker vi her om et paradigmeskifte i institusjonsbibliotekene? Behovene for bibliotek-tjenester endres og utvides. Tidligere var det kun snakk om å hente inn ekstern informasjon og å gjøre denne tilgjengelig for institusjonens brukere. Nå er det også behov for å gjøre institusjonens materiale tilgjengelig for eksterne brukere (Jones, Andrew, MacColl 2006, preface xv).

I Bibliotekreform 2014 har den raske teknologiske utviklingen og tilgangen til digitalt materiale fått stort fokus. Man sier at fornyelse av bibliotekenes tjenester i stor grad vil være knyttet til at bibliotekene kan tilby brukervennlige og lett tilgjengelige digitale og nettbaserte tjenester. Et av hovedelementene i bibliotekreformen er satsningen på digitale tjenester og lett tilgjengelig nettbasert innhold.

*”Utviklingen av åpne institusjonsarkiv, som er en internasjonal satsning i universitets- og høyskolesektoren, er et viktig framskritt for formidling av forskningslitteraturen og bidrar til at den blir en del av den digitale kunnskapsallmenningen.*

...

*Bibliotekene skal være en viktig kanal for tilgang til kunnskaps- og kulturressurser og utgjøre en digital kunnskapsallmenning. For å gjøre dette mulig må Norgesbiblioteket<sup>1</sup> satse på utvikling av nettbaserte tjenester både i et samarbeid mellom bibliotekene og i hele abm-sektoren<sup>2</sup>.*

*Digitale tjenester som legges åpent ut på Internett er i prinsippet tilgjengelige for alle. ... Ved å eksponere metadata for søkemotorene sikres synlighet og tilgjengelighet på flere måter.”*

(ABM-utvikling 2006, s.16-17).

---

<sup>1</sup> **Norgesbiblioteket** – betegnelse for et samarbeidende nettverk av alle bibliotektyper. Det omfatter alle offentlige bibliotek og alle medietyper, det omfatter både konvensjonelle og nettbaserte tjenester (ABM-utvikling 2009).

<sup>2</sup> **ABM-sektoren** – samlebetegnelse for Arkiv, bibliotek og museum.

Tilgjengelighet og åpenhet er flott, men det stiller krav til gjenfinning. Jeg er opptatt av hva som skjer med alt som produseres både i det private og i det offentlige. Hvor blir det av? Legges det i hyller hvor det samler støv og blir glemt, eller blir produksjonen publisert og tilgjengeliggjort? Tilgjengeliggjort for hvem, på hvilken måte og blir det funnet igjen? Og, ikke minst, blir det brukt? Dette berører etter min mening kjernen ved ideen om Open Access publisering (OA) og institusjonelle arkiv (IA).

Med bakgrunn i dette har jeg gjennomført en studie hvor jeg har sett på publisering og tilgjengeliggjøring av vitenskaplig og faglig materiale i den norske universitets- og høgskolesektoren. Spesifikt har jeg sett på innføring og bruk av institusjonelt arkiv (IA) som ny kanal for dette. Hovedformålet har vært å undersøke om innføringen av IA har hatt en positiv effekt på gjenfinning og bruk av institusjonens forskningsproduksjon – gir IA en verdiøkning? Er IA et verktøy som utgjør en forskjell for akademiske institusjoners samlinger, et gode for synliggjøring av forskningen og institusjonen?

## 1.1 Open Access

Ny teknologi og Internett har gitt helt nye muligheter for spredning av vitenskaplige og faglige arbeider gjennom verdensomspennende elektronisk distribusjon. Definisjonen av Open Access er blitt raffinert etter hvert som ”bevegelsen” har utviklet og bredt om seg, men fortsatt er definisjonen fra Budapest Open Access Initiative i bruk. Møtet i Budapest i regi av Open Society Institute, som resulterte i initiativet, er betegnet som den offisielle starten på Open Access bevegelsen (OSI 2002).

*”... The literature that should be freely accessible online is that which scholars give to the world without expectation of payment. Primarily, this category encompasses their peer-reviewed journal articles, but it also includes any unreviewed preprints that they might wish to put online for comment or to alert colleagues to important research findings. ... By "open access" to this literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself.” (OSI 2002, 3.avsn.)*

Open Access betyr at fagfelleverderte<sup>3</sup> vitenskapelige artikler og annet faglig materiale gjøres åpen og fritt tilgjengelig via Internett (online og gratis for allmennheten). OA brukes som virkemiddel for å sikre større tilgjengelighet til, og bruk av en institusjons forskningsresultater. Forfatter eller opphavsmann beholder opphavsretten til publikasjonen, men gir brukere tillatelse til å lese, laste ned, kopiere, distribuere, skrive ut, søke i eller lenke til fullteksten uten å forlange vederlag.

Det er definert 2 alternative veier til Open Access:

1. *Parallellpublisering/egenarkivering* - forfatteren arkiverer sitt materiale, for eksempel en artikkel, i institusjonens åpne og fritt tilgjengelige *institusjonelle arkiv*, hvor artikkelen vil være tilgjengelig i fulltekst. I tillegg blir artikkelen publisert i tradisjonelle tidsskrifter. Det vil si at det foretas en parallellpublisering i IA basert på avtaler om rettigheter. Denne strategien kalles ”the green rOAd to open access”.
2. *Open Access tidsskrifter* - forfatteren publiserer sine artikler i OA tidsskrifter, dvs. at forlaget bak OA tidsskriftene velger å la tidsskriftene være åpne og fritt tilgjengelige på internett. Denne strategien kalles ”the golden rOAd to open access”.

(Harnard 2008)

**Institusjonelt arkiv** – en åpen og fri digital samling av den vitenskapelige produksjonen ved en forsknings- og/eller utdanningsinstitusjon. Et IA inneholder alle typer faglige og/eller vitenskaplige materiale – alt fra mastergradsoppgaver, doktorgradsavhandlinger, rapporter, tidsskriftartikler, bøker, bokkapitler og konferansepapirer til kunstnerisk materiale som tegninger, fotografier og musikk, læringsobjekter, posters og til og med rådata og datasett. Et IA kan omfatte materiale fra alle avdelinger/fakulteter ved en institusjon og kan vise institusjonens samlede produksjon. Det kan også være et fagbasert arkiv, dvs. at det kan omfatte et enkelt fag eller avgrensede fagavdelinger ved institusjonen. En studie som undersøkte forlagenes tillatelse til parallellpublisering fant ut at 47 % av alle norske vitenskaplige artikler produsert i perioden 2005-2006 kunne vært parallellpublisert/egenarkivert i IA (Hernes 2008).

---

<sup>3</sup> **Fagfellevurdering (peer review)** – gjennomgang og vurdering av et vitenskapelig arbeid som er en viktig kvalitetssikrende faktor i vitenskapelig publiseringen. Metodens styrke ligger i at fagpersoner er mest kompetente til å avgjøre betydningen og verdien av et manuskript eller å oppdage opplagte metodiske eller logiske feil.

**Open Access tidsskrift** – et tidsskrift uten abonnementsavgift, dvs. gratis for sluttbrukeren. Kostnadene ved å drive slike tidsskrift dekkes via et publiseringsgebyr som betales av forfatteren selv, forfatterens finansieringskilde eller institusjonen forfatteren er ansatt ved. Betales gebyret av institusjonen vil kostnaden bli regnet som en del av budsjettet for forskningsprosjektet, eller som del av daglig drift. Directory of Open Access Journals er et nettsted som gir en oversikt over Open Access tidsskrifter (DOAJ 2008). Pr april 2009 besto oversikten av i alt 4000 tidsskrifter.

## **1.2 Internasjonal og nasjonal status**

En rekke organisasjoner innen vitenskaplig forskning har de senere år lagt fram henstillinger om mer åpenhet. Open Access bevegelsen sprer seg både internasjonalt og nasjonalt, og i det følgende har jeg gitt en oversikt over noen organisasjoner og prosjekter som har hatt, og har sterk innflytelse. Initiativene listes i kronologisk rekkefølge, først internasjonale tiltak, og deretter tiltak og status i Norge.

### **Open Society Institute**

Open Society Institute (OSI) ble stiftet i 1993 av investoren og filantropen George Soros. OSI er en privat stiftelse som jobber med utforming av offentlig politikk og strategi for å fremme demokratiske styresett, menneskerettigheter og økonomiske, rettslige og sosiale reformer på verdensbasis. OSI jobber med flere OA prosjekter, og står blant annet bak Budapest Open Access Initiativet. De har påtatt seg ansvaret for, og bruker ressurser på, å utvikle og fremme dette initiativet videre. I tillegg støtter de lanseringen av nye OA tidsskrifter og er rådgivere for å få disse økonomisk selvstendige og bærekraftige (OSI, 2008).

### **Budapest Open Access Initiative**

Open Society Institute arrangerte et møte i Budapest i slutten av 2001 hvor agendaen var forandringer i vitenskaplig kommunikasjon. Resultatet av dette møtet blir kalt Budapest Open Access Initiativ, som er en uttalelse om prinsipper, strategier og forpliktelser i forhold til å arbeide for åpen tilgang til vitenskapelig materiale. Både OA publisering og IA, henholdsvis ”the golden rOAd”- og ”the green rOAd”-strategiene, ble lansert i Budapest Open Access Initiative. Som nevnt tidligere betegnes initiativet som den offisielle starten på OA bevegelsen. Pr april 2009 hadde 489 organisasjoner og 5009 enkeltindivider signert Budapest Open Access Initiative (OSI 2002).

### **Bethesda Statement on Open Access Publishing**

På slutten av 2002 ble det avholdt et nytt møte om Open Access på Howard Hughes Medical Institute i Chevy Chase, Maryland, USA. Resultatet ble også her en felles uttalelse – Bethesda Statement on Open Access Publishing. Uttalelsen baserer seg på Budapest Open Access Initiativ, men utvider OA begrepet ved å si noe om hvordan forfatterne eller rettighetshaverne skal operasjonalisere/utøve OA konseptet. En brukers rettigheter reguleres gjennom lisenser, såkalte Creative Common Licenses, og blant annet oppfordres det til å publisere en brukers rettigheter i tillegg til det vitenskaplige materialet (Howard Hughes Medical Institute 2003).

### **Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities**

Et kanskje enda mer kjent dokument er Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, kaldt Berlindeklarasjonen. Dette er en oppfølger av både Budapest Open Access Initiative og Bethesda Statement on Open Access Publishing. Ordlyden er mye av den samme som i Bethesda Statement. Deklarasjonen var pr april 2009 signert av 262 organisasjoner (Berlindeklarasjonen 2003).

### **European Research Council**

European Research Council kom med sin uttalelse om Open Access i desember 2006, hvor de anbefaler at forskere og institusjoner går sammen om åpen og fri tilgang (ERC 2006).

### **European Research Advisory BOArd**

Anbefalingen er fulgt opp i rapporten “Scientific Publication: Policy on Open Access”. Her gir European Research Advisory BOArd sin anbefaling til Europakommisjonen angående EUs 7. rammeprogram, hvor de sier at alle forskningsresultater knyttet til forskning som er støttet av FP7 (The Seventh Framework Programme) bør parallellpubliseres i IA så raskt som mulig, og senest innen 6 måneder, etter publisering (EURAB 2008).

### **Harvard University**

Som et av de første vedtok fakultetet for humaniora og naturvitenskap ved det amerikanske universitetet Harvard i februar 2008 at alle ansatte skal gjøre sine forskningsarbeider åpne og fritt tilgjengelige gjennom et institusjonelt arkiv. I mai 2008 fulgte det juridiske fakultetet opp:



*“On May 1, 2008, the Harvard Law School faculty voted to adopt an open access policy making their scholarly articles freely available to the public. Under the policy, faculty members negotiating contracts with publishers retain copyright to their scholarly articles and grant the fellows of Harvard College permission to make them available in Harvard's open access repository.”*

(Harvard University 2008).

### **Universitets- og høskolerådet**

Universitets- og høskolerådet (UHR) kom i februar 2005 med en henstilling til sine medlemsinstitusjoner om å ta i bruk åpne institusjonelle arkiv og publisering i Open Access tidsskrifter, og 30.01.09 sendte de en rapport til Kunnskapsdepartementet som svar på en henvendelse om råd i forhold til å stimulere til mer åpen tilgang til forskningsresultater. UHR mener at tilgang til vitenskapelige resultater er et viktig forskningspolitisk spørsmål, og at en positiv holdning til OA initiativet fra Kunnskapsdepartementets sin side vil få stor betydning for utviklingen ved institusjonene. UHR sier i sin rapport at Kunnskapsdepartementet bør støtte åpen tilgang gjennom utforming av politiske mål og i sin oppfølging av universiteter og høyskoler. Videre bør Kunnskapsdepartementet oppfordre institusjonene til å utarbeide egne strategier og retningslinjer som bidrar til åpen tilgang til vitenskapelige publikasjoner (UHR 2005, UHR 2009).

### **Norges forskningsråd**

Norges forskningsråd vedtok i februar 2009 nye prinsipper for åpen tilgang til vitenskapelig publisering. Disse slår fast at vitenskapelige tidsskriftsartikler som bygger på FoU-prosjekter finansiert av Forskningsrådet, skal lagres i åpne digitale arkiv og dermed gjøres tilgjengelig for alle interesserte. Det understrekes imidlertid at slik arkivering ikke må bryte med forfatteres og utgiveres rettigheter (Forskningsrådet 2009a).

### **Det norske fag- og forskningsmiljøet**

Holdningen til Open Access er positiv også i det øvrige norske fagmiljøet. Flere større finansieringskilder for forskning jobber nå med strategier hvor de anbefaler at forskningsresultatene enten publiseres i OA tidsskrifter eller parallellpubliseres i IA. Policybeslutninger om krav til publisering er også på trappene. Mange mener at om ikke annet, bør i det minste offentlig finansiert forskning være åpen og offentlig tilgjengelig. Når forskningsresultater publiseres åpent blir resultatene tilgjengelig for alle forskningsmiljøene – også de som ikke har mulighet til å betale for dette.

Mye er allerede fulgt opp. Institusjonelle arkiv er implementert ved alle landets universiteter, og er, eller er i ferd med å bli implementert også ved høyskolene. Retningslinjene divergerer noe fra institusjon til institusjon. De fleste har ikke innført krav, men en sterk henstilling om arkivering. Universitetene i Oslo og Trondheim stiller i dag krav om obligatorisk innlevering og arkivering av mastergradsoppgaver og doktorgradsavhandlinger i elektronisk versjon, men fortsatt gir de kun en henstilling om innlevering og arkivering av annet faglig og vitenskaplig arbeid (Jakobsson 2003, NTNU 2008).

Til tross for anbefalinger og gode hensikter knyttet til muligheten for økt anerkjennelse gjennom flere siteringer, er artikkelforfatterne/forskerne ikke like begeistret for ideen om publisering i OA tidsskrifter eller parallellpublisering i IA som ledelsen ved forskningsinstitusjonene. Mange er skeptiske til forvanskning av opphavsrettslige forhold, endring av betalingsstruktur og økte administrative oppgaver for den enkelte. Spørsmålet som stilles er "What's in it for me"? Etter offentliggjøringen av anbefalingene til Kunnskapsdepartementet fra både UHR og Forskningsrådet tidligere i vinter har det vært skrevet flere avisartikler som beskriver ulike synspunkter. I en artikkel i Aftenposten (Bjørkeng 2009) uttrykker forsker og paleontolog Jørn Hurum at anbefalingene er et *"kjempebra fremstøt"*. Han sier følgende: *"Det er jo helt dust at vi gir fra oss rettighetene til forlagene", og "Universiteter og Forskningsrådet betaler lønn både til de som skriver, leser og gjerne også til redaktørene i de mindre tidsskriftene. At resultatene da ikke er offentlig tilgjengelig er så dumt og gammeldags som det går an"*. I den samme artikkelen uttrykker forhandlingssjef Frank O. Anthun i Forskerforbundets opphavsrettsutvalg at han synes det er betenkelig at betalingen flyttes fra leser til forsker gjennom åpen publisering, dvs. publisering i OA tidsskrifter, og hevder at det bare er de miljøene som har økonomi som i et slikt regime kan betale seg til publisering. Når det gjelder egenarkivering, eller parallellpublisering i IA, mener han det er opp til den enkelte forsker – *"det handler om den akademiske frihet"*. Trond Andreassen i Norsk faglitterær forfatter- og oversetterforening, som organiserer mange av forskerne og forvalter deres rettigheter, er også skeptisk. Hans kommentar er: *"Vi er spørrende og avventende negative til disse nye kravene fra Forskningsrådet"*.

Jeg skal ikke ta opp diskusjonen om forskernes holdninger videre her i denne oppgaven, det er det andre som skriver om, men min undersøkelse tar utgangspunkt i forskerne og deres produksjon, slik at resultatene mine kanskje kan bidra til å gi noen nye argumenter i denne diskusjonen.

### 1.3 Institusjonelle arkiv - formål

Formålet med et IA er synliggjøring av institusjonen. Arkivet blir institusjonens ansikt utad i forhold til å tilby en samlet presentasjon av institusjonens egenproduserte publikasjoner – alt digitalt, åpent og fritt tilgjengelig via nett. Innføring av IA er begrunnet blant annet i følgende:

1. Synliggjøring av institusjonen
2. Presentasjon og tilgjengeliggjøring av institusjonens egen forskning
3. Tilgjengeliggjøring av materiale som tradisjonelt ikke publiseres
4. Bevaring
5. Støtte til Open Access bevegelsen

Punktene presenteres og utdypes mer detaljert i fortsettelsen.

#### 1. IA er både et hensiktsmessig og nyttig instrument for synliggjøring av institusjonen.

- Det er nyttig for institusjonen å kunne arkivere/tilgjengeliggjøre all produksjon ved institusjonen; studentoppgaver som bachelor- og mastergradsoppgaver, doktorgradsavhandlinger, faglige og vitenskaplige artikler og rapporter (pre- og postprints<sup>4</sup>), konferansepapirer, lokale skriftserier og annet materiale produsert ved institusjonen.
- Det er hensiktsmessig og nyttig for studenter, forskere og ansatte ved en institusjon å ha muligheten til å søke i og bruke en lett tilgjengelig database/et felles arkiv som inneholder institusjonens forskning. Faglig og profesjonelt materiale som kan inkludere alt fra utkast til fagfelleverderte artikler, bilder, audio, video og annet kan være verdifull for andre i deres forskning.
- Det er nyttig fordi det gir oversikt over hvilke prosjekter og hvilken forskning som gjennomføres ved institusjonen på tvers av fakulteter og avdelinger. Dette åpner for samarbeid og videre fagfellevurdering.

---

<sup>4</sup> **Preprint** refererer til en artikkel **før** fagfellevurdering, eller sagt på en annen måte; en artikkel som ikke er akseptert og publisert i et vitenskaplig tidsskrift enda.

**Postprint** refererer til den bearbejdede artikkelen **etter** fagfellevurdering, dvs. en artikkel som er akseptert og kan publiseres eller er publisert i et vitenskaplig tidsskrift.

- Det er hensiktsmessig som markedsføring av institusjonens fakulteter og avdelinger overfor mulige fremtidige studenter eller forskere som ønsker å søke seg til institusjonen. Også nåværende studenter som søker etter en ekspert innen det tema han/hun er interessert i vil kunne finne nyttig informasjon i institusjonens IA.
2. IA er en hensiktsmessig kanal for å presentere/tilgjengeliggjøre institusjonens egen forskning for forskningsmiljø og andre interessenter utenfor institusjonen. Dersom IA støtter OAI-PMH protokollen<sup>5</sup>, kan metadata<sup>6</sup> høstes og gjøres tilgjengelige både i nasjonale og internasjonale søketjenester og søkemotorer. Tilgjengeliggjøringen har potensial for å gi akademisk anerkjennelse, øke institusjonens prestisje og gi større forskningsbetydning – research impact – gjennom økt sitering.
  3. IA er en hensiktsmessig kanal for å tilgjengeliggjøre materiale som tradisjonelt ikke publiseres, for eksempel studentoppgaver som bachelor- og mastergradsoppgaver og doktorgradsavhandlinger. Fordi disse dokumenttypene vanligvis ikke publiseres offentlig har slike dokumenter kun vært tilgjengelig ved søk i institusjonens bibliotekskatalog og gjennom ordinært lån/fjernlån. Gjennom parallellpublisering i IA kan materiale nå synliggjøres og gjøres lett tilgjengelig i en elektronisk form for miljøer både innenfor og utenfor institusjonen.
  4. IA er et nyttig instrument for bevaring. Flere og flere institusjoner går fra rent papirbaserte arkiver til elektroniske arkiver, og spørsmålet om bevaring for fremtiden uavhengig av formater og lagringsmedier står sentralt. Gjennom bruk av IA kan ansvaret for bevaring flyttes fra den enkelte forfatter (student, forsker, ansatt) til institusjonen. Strategier, rutiner og ikke minst ressurser for fremtidig bevaring får fokus og blir ivaretatt i en større sammenheng.
  5. IA er et nyttig instrument i kampen mot en stadig økning i abonnementsprisen på tidsskrifter og databaser hvor vitenskaplige og faglige arbeider blir publisert. Ved å gjøre

---

<sup>5</sup> **OAI-PMH protokoll** - Open Archive Initiative Protocol for Metadata Harvesting. Åpen standard for utveksling av metadata. Standarden legger ikke opp til å hente en og en post, men grupper av poster som høsting med ”skurtresker”. ”Latterskel-protokoll” som er lett å implementere hos IA-institusjonene.

<sup>6</sup> **Metadata** – data om ressursen/dokumentet, dvs. data som beskriver ressursen, for eksempel tittel, forfatter, utgiver, publiseringsdato, etc. Disse dataene kodes etter et valgt metadataskjema og den mest kjente og brukte er Dublin Core.

majoriteten av egen produksjon åpent tilgjengelig via et IA med minimale restriksjoner for videre bruk støtter institusjonen Open Access bevegelsen. Et IA gir mulighet for global tilgjengeliggjøring, dvs. tilgjengeliggjøring både for miljø som i utgangspunktet har ressurser til selv å abonnere på dyre tidsskrifter, men også for miljø som ikke har disse ressursene.

(Buehler & Boateng 2005).

## 2. Litteratur og tidligere forskning

Det er skrevet mye om Open Access generelt og om institusjonelle arkiv og Open Access publisering spesielt. Jeg startet å søke helt åpent på nettet ved bruk av søkebegrepene ”institusjonelle arkiv” og ”Open Access publisering” i ulike institusjonelle arkiv i Norge, i bibliotekataloger, i Open Access tidsskrifter og ulike internasjonale søkeportaler/innhøstere. Det resulterte i kolossale mengder med informasjon, og for å gi en forhåpentligvis bedre oversikt har jeg satt opp 3 kategorier av materiale tilknyttet temaet.

### 2.1 Institusjonelle arkiv i utlandet

Det finnes mange artikler og rapporter fra universiteter i utlandet om deres erfaringer med implementering av IA. Rapporter både fra USA og England underbygger begrunnelsen for innføring av IA og støtte til OA bevegelsen.

I tillegg til de faglige fordelene som trekkes frem ved bruk av Open Access og IA, beskrevet i kapittel 1.3, blir også bruk av åpenhet og standarder sett på som fordeler for blant annet drift av IA. Bruk av open source programvare spesielt vinklet på IA gjør verden enklere for en driftsansvarlig, som oftest en bibliotekar. Her kan jeg nevne systemløsninger som DSpace, Fedora og EPrint, bruk av felles metadatastandarder basert på Dublin Core, bruk av OAI-PMH protokoll som gir mulighet for høsting, opprettelsen av ulike åpne registre som gir oversikt over OA tidsskrifter og hvilke tidsskrifter som tillater parallellpublisering/egenarkivering, etc. (Barwick 2007).

Eksempler på åpne registre er:

- SHERPA/RoMEO (Publisher copyright policies & self-archiving): gir oversikt over hvilke rettigheter et tidsskriftsforlag gir sine artikkelforfattere. Rettighetene legges

ofte inn som en del av copyright transportavtalen, som benyttes ved publisering i fagtidsskrifter (SHERPA/RoMEO 2008).

- OpenDOAR (Directory of Open Access Repositories): gir oversikt over kvalitetssikrede IAer (OpenDOAR 2008).
- DOAJ (Directory of Open Access Journals): gir oversikt over kvalitetskontrollerte faglige og vitenskaplige Open Access tidsskrifter og tilbyr artikler i fulltekst (DOAJ 2008).

Mange er positive, men ikke alle. En av dem som stiller spørsmålstegn er Dorothea Salo (2007), som er bibliotekar ved Universitetet i Wisconsin, USA, og ansvarlig for universitetets IA. Hun peker på problemene ved at forskerne ikke ser de samme fordelene som bibliotekarene. Det går veldig tregt med parallellpublisering/egenarkivering, og det må pushes og lokkes. Holdningen blant forskerne er fortsatt at publisering i fagfelleverderte tidsskrifter er viktigst. Gjennom disse får forskningen og forskeren betydning – ”impact”, ved at artiklene leses og er tilgjengelige for forskerne innen samme fagfelt. Salo tar også opp problemet med at IAer driftes av bibliotekarer som ofte ikke har nok faglig dyktighet til å kvalitetssikre det som egenarkiveres. Kvaliteten på metadataene blir ikke god nok, og feil forekommer for hyppig.

## 2.2 Institusjonelle arkiv i Norge

Det finnes noen rapporter, om ikke så mange, fra Norge som beskriver fenomenet sett med norske øyne. Jeg har lest om og prøvd/testet flere IAer hos universiteter og høyskoler i Norge, samt NORA (Norwegian Open Research Archives), som er den nasjonale søketjenesten for åpne institusjonelle arkiv. NORA er en innhøster, dvs. tjenesten høster/henter metadata fra arkiv rundt omkring i Norge for å tilby en felles søketjeneste. NORA ble opprettet som et prosjekt for perioden 2005-2007, men ble videreført i 2008 og nå i 2009 basert på finansiering fra Kunnskapsdepartementet. I NORAs prosjektbeskrivelse skrives det følgende:

*”...Den nasjonale søketjenesten skal bruke OAI-PMH for å samle inn metadata fra de lokale åpne institusjonelle arkivene og gjøre dem søke- og presenterbare i ett enhetlig grensesnitt. Når antallet åpne institusjonelle arkiv øker, øker behovet for interoperabilitet. Man trenger et konsistent miljø for utbytte, spredning og gjenvinning av metadata. Å samordne disse skaper*

*en forutsetning for meningsfulle tjenester på forskjellige nivåer, både nasjonalt og internasjonalt. Høstingen forenkles hvis de åpne institusjonelle arkivenes modeller er harmonisert ved hjelp av en felles metadatamodell. Den felles metadatamodellen utgjør grunnen for forbedrede søke- og presentasjonsmuligheter innen den nasjonale søketjenesten. Det er en klar fordel for de lokale åpne institusjonelle arkivene å eksponeres i en nasjonal søketjeneste, da det gir maksimal synlighet også for små arkiv.” (UiT 2008a)*

Arbeidsgrupper ved flere universitet og høyskoler i Norge har utarbeidet forprosjektrapporter med anbefalinger om innføring av IA. Rapportene belyser hva institusjonene fokuserer på og hvilke problemstillinger de står ovenfor ved innføring av et IA. Flere av de fordelene og utfordringene som er beskrevet tidligere i oppgaven blir også diskutert i disse rapportene. (Sølvberg 2007, Fagerjord 2007 og Pharo 2008).

I forbindelse med markeringen av 14.oktober 2008, som var verdens første Open Access dag, skriver Universitetsbiblioteket i Oslo følgende om DUO, som er UiOs eget IA:

*”DUO blir større for hver dag som går, både i antall dokumenter og antall nedlastninger. Ved siste opptelling var det totalt 8493 fulltekstdokumenter i DUO, av disse var 278 fagfelleverderte. Hittil i år har det vært over 160000 nedlastninger i DUO. Det mest populære dokumentet har vært lastet ned hele 1135 ganger.” (UiO 2008).*

For å sette dette inn i en sammenheng er beholdningen i DUO pr juni 2009 totalt på 8978 dokumenter, og av disse er 404 fagfelleverderte. Det betyr at andelen av fagfelleverderte dokumenter i dag utgjør 4,5 % av dokumentene totalt mot 3,3 % i oktober 2008. Det går ikke raskt, men likevel fremover med den vitenskaplige parallellpubliseringen/egenarkiveringen hos UiO.

### **2.3 Institusjonelle arkiv og ”impact”**

Selv om interessen er stor og ideen brer om seg, så er fagfeltet Open Access forholdsvis nytt. Det er derfor ikke gjennomført mange vitenskaplige undersøkelser og forskningsprosjekter på feltet, men noe er det. Et fellestrekk er at de baserer seg på, og refererer til de samme få personene som har gjennomført tidligere arbeider på feltet; Steve Lawrence, Kristin Antelman, Stevan Harnad, Tim Brody, Michael Kurtz, Steve Hitchcock og noen til. Et annet

fellestrekk er at alle analysene beregner betydningen av vitenskaplig arbeid (research impact) i antall siteringer, og i noen få tilfeller i antall nedlastinger.

En av de første som gjennomførte en access/impact-analyse var Steve Lawrence (2001). Han fant ut gjennom å sammenligne antall siteringer på artikler publisert som OA artikler og som ikke-OA artikler at i snitt ble OA artikler sitert 2.56 ganger mer sammenlignet med ikke-OA artikler, dvs. forskjellen var på hele 156 %. Han analyserte ca. 120.000 konferanseartikler innen vitenskapsdisiplinen datateknologi og relaterte disipliner produsert over en periode på 10 år. Han brukte ResearchIndex, som er en siteringstjeneste hos IST (Penn State's College of Information Sciences and Technology), for å estimere antall siteringer og online tilgjengelighet. Lawrence sier selv at datamaterialet han brukte ikke var godt nok til å kunne gi et vitenskaplig svar, men at tallene ga indikasjon på at bedre tilgjengelighet og bedre synlighet på OA artiklene spilte en signifikant rolle for betydningen – impact'en.

Kristin Antelman bekrefter Lawrences teorier i sitt forskningsarbeid (2004). Hennes hypotese den gang var at vitenskaplige artikler fra vitenskapsdisipliner med ulik tilnærming til Open Access hadde større innflytelse og betydning (research impact) dersom artiklene var åpent tilgjengelige enn om de ikke var det. Hun plukket ut artikler fra de 10 høyest rangerte tidsskriftene for 4 ulike vitenskapsdisipliner fra ISI's<sup>7</sup> Journal Citation Reports<sup>8</sup> og brukte ISI Web of Science<sup>9</sup> for å finne antall siteringer (les: artiklenes impact). Videre brukte hun Google for å finne ut om artiklene i tillegg var publisert åpent. Hun kom frem til at det var en signifikant forskjell på antall siteringer i snitt på OA artikler sammenlignet med ikke-OA artikler. Resultatet var en 45 % økning på antall siteringer innen filosofi, 51 % innen elektronikk og ingeniørfag, 86 % innen politisk forskning og hele 91 % innen matematikk.

Stevan Harnad og Tim Brody bekrefter dette på ny gjennom nye undersøkelser senere i 2004 (Brody & Harnad 2004, Harnad et al. 2004). Disse undersøkelsene tar utgangspunkt i artikler fra mange ulike vitenskapsdisipliner publisert i tidsskrifter som tillater parallellpublisering/

---

<sup>7</sup> **The Institute for Scientific Information** (ISI) ble grunnlagt i 1960 av Eugen Garfield. De har siden starten tilbudt bibliografiske databasetjenester med spesielle innen siteringsindeksing, -søk og -analyse. ISI ble i 1992 overtatt av Thomson Scientific & Healthcare, ble kjent som Thomson ISI, og har senere fått navnet Thomson Scientific (Thomson Reuters 2009).

<sup>8</sup> **Journal Citation Reports** (JCR) er en årlig publisert liste som gir oversikt over "impact factor" for de journalene som er med i undersøkelsen. Rangeringen er basert på et sett med offentlig publiserte kriterier.

<sup>9</sup> **Web of Science** består av syv databaser som inneholder informasjon fra over 10.000 vitenskaplige tidsskrifter, bøker, serier, rapporter og proceedings fra internasjonale konferanser, symposier, seminarer, etc., som er samlet for ca 250 vitenskapsdisipliner.



egenarkivering. De sjekket siteringsraten på artiklene og så dette i forhold til om de i tillegg var egenarkivert eller ikke. Den største av undersøkelsene omfattet opptelling av antall siteringer på 14 mill artikler på tvers av alle vitenskapsdisipliner over en 11-års periode. Som Antelman benyttet også Harnad og Brody ISI Web of Science og Science Citation Index. Sjekken av parallellpublisering ble gjort ved å sjekke hvilke av artiklene som ble parallellpublisert i IA samme år de ble tidskriftspubliseret. Her ble dels arXiv (Cornell University Library 2008) og dels andre søketjenester benyttet. Antall OA artikler utgjorde i perioden 1992-2003 fortsatt en liten andel av artiklene publisert totalt, men siteringsraten på OA artikler lå langt over ikke-OA artikler. For disiplinen fysikk/matematikk var andelen av OA artikler i 1992 kun 2 % av antall artikler totalt, men disse hadde 87 % høyere siteringsrate enn ikke-OA artikler. I 2003 var tallene steget til henholdsvis 10 % andel artikler og 79 % høyere siteringsrate. Det bør presiseres at antall artikler totalt tilgjengelig, dvs. både OA/ikke-OA, hadde økt fra ca 5000 i 1992 til nesten 1.6 mill i 2003.

Konklusjonen fra tidligere forskning er altså helt entydig på at bruk av Open Access med åpen og fri tilgjengelighet til faglige og vitenskaplige forskningsresultater gir økt lesing og økt antall sitering, derigjennom økt betydning og innflytelse – større ”impact”. Sagt med andre ord; ”It’s something in it for me”.

Norge har ikke kommet like langt når det gjelder Open Access, men de siste par årene har det vært en eksplosjon i antall IAer. Pr april 2009 er materiale fra 31 utdanningsinstitusjoner og 8 forskningsbaserte institutter og direktorater – totalt 39 institusjoner, tilgjengelig via NORA. Ifølge NOKUT (Nasjonalt organ for kvalitet i utdanningen) er det 76 institusjoner som betegnes som utdanningsinstitusjoner, og som alle driver forskning i større eller mindre grad i tillegg til undervisning. Når kun 31 av disse har innført IA, og 10 av disse igjen har hatt oppstart det siste halve året, betyr det at det fortsatt er en lang vei å gå. Antall dokumenter har økt med 10,7 % fra oktober 2008 til april 2009, men med sine 22.067 dokumenter totalt er likevel NORA ”lillebror” i forhold til den svenske søketjenesten DiVA som pr april 2009 hadde 170.549 forskningsdokumenter. I tillegg inneholder DiVA 32.462 studentoppgaver (UiO 2009 & Uppsala universitet 2009).

## **2.4 Institusjonelle arkiv i et strategisk og forvaltningsmessig perspektiv**

Forskningstemaet i mitt prosjekt er knyttet til graden av tilgjengelighet av faglig og vitenskaplig materiale opp mot graden av hvilken betydning og innflytelse – ”impact”,

tilgjengeligheten gir. Når jeg skal sette dette inn i et teoretisk perspektiv velger jeg å knytte det opp mot tanker ulike forskningsmiljøer har for vitenskaplig kommunikasjon i Norge.

Forskningsrådets visjon:

*”Forskning og utvikling utvider grensene for hva vi vet, forstår og kan få til. Det gjør samfunnet kulturelt rikere og gir samtidig grunnlag for en bærekraftig utvikling, økt velferd og økt verdiskaping.*

*Det kunnskapsbaserte samfunnet krever at forskningen er i fronten internasjonalt. Forskningsrådet vil styrke norsk forsknings internasjonale posisjon gjennom å være en endringsagent i samvirke med aktørene i forskningssystemet. Forskningsrådet skal fremme forskning av høy kvalitet og bidra til at forskning anerkjennes som en kulturell og verdiskapende kraft i utviklingen av samfunnet.”*

(Forskningsrådet 2009b, s. 3)

Forskningsrådets strategidokument for perioden 2009-2012 ”I front for forskning” har satt opp 4 utfordringer som viser hva Forskningsrådet fokuserer på for å følge opp myndighetenes forskningspolitikk i årene fremover. Disse er (utheving gjort av meg):

- Kapasitets- og kvalitetsutfordringen – hvor man er opptatt av at innsatsen og kvaliteten i forskningen må økes og bli bedre. Kvalitetsbegrepet skal vurderes ut fra et disiplinbasert perspektiv, men også i **samspill mellom forskning og anvendelse**.
- Samfunnsutfordringen – hvor man er opptatt av at forskningen i større grad må svare på spesifikke utfordringer i samfunn og næringsliv. Forskning må gi læring og nytte gjennom at **forskningens resultater må formidles bedre** og evnen til å utnytte resultatene må forbedres. I så henseende vil **gode lærings- og innovasjonsarenaer** legger til rette for en bred utnyttelse av forskningsresultatene.
- Strukturutfordringen – hvor man er opptatt av **forskingssystemets struktur, samvirke og internasjonale deltakelse**. Her snakkes det om samarbeid og sterke samlinger av miljøer/ institutter/regioner/næringsklynger av utvalgte satsningsområder, og viktigheten av økt internasjonalisert forskningssamarbeid.

- Læringsutfordringen – hvor man er opptatt av at forskningen må skape resultater som er nyttige for næringsliv og forvaltning, og den må gi læring som styrker kunnskapskulturen i samfunnet. Her trekkes det spesielt frem at det er **stor åpenhet om resultatene** og at den **allmennrettete formidlingen** må styrkes.

Forskningsrådet er av den oppfatning at forskning blir stadig viktigere som både innovasjonskraft, problemløser, dannelses- og kunnskapsprodusent, og at den globale konkurransesituasjonen gjør at kunnskapsutviklingen akselererer, og kunnskap blir nasjonens viktigste konkurransefortrinn (Forskningsrådet 2009b).

Universitetet i Oslo har ett av Norges største forskningsmiljø. I sin strategiske plan for 2005-2009 er et av hovedmålene at *”UiO vil bidra aktivt til at forskningsbasert kunnskap nyttiggjøres i samfunnet og arbeide for å gi allmennheten innsikt i vitenskapelig tenkning og oppdatert kunnskap om forskningens resultater”*. For å nå dette hovedmålet vil UiO bl.a. fremme kvalitet og gjennomslag i allmennrettet formidling og gi bedre tilgangen til UiOs kunnskapskilder ved at *”... ”Fjerntilgangen” til UiOs samlinger og publikasjoner skal bedres ytterligere gjennom videre digitalisering av publikasjoner og samlinger og tilrettelegging av elektroniske tjenester. Museenes og bibliotekets samlinger og kompetanse skal nyttiggjøres bedre i øvrig formidlings- og kontaktvirksomhet.”* (UiO 2005, s.11)

Openaccess.no er en norsk informasjonstjeneste for alle som trenger informasjon om, og som vil utveksle tips omkring temaet Open Access. Redaksjonen er lokalisert hos Universitetsbiblioteket i Tromsø. De sier følgende:

*”Hensikten med Openaccess.no er å bidra til effektive system for åpen vitenskapelig kommunikasjon på linje med forskningens grunnleggende mål: - bredest mulig spredning av forskningsresultater. Openaccess.no skal bidra til at norske forskningsresultater blir synlige og åpent tilgjengelige, uten begrensinger, for alle brukere nasjonalt og internasjonalt.”*  
(UiT 2008b)

Open Access er altså et av virkemidlene som skal bidra til å nå målene om åpenhet og tilgjengelighet, og skal bidra til å gi vitenskapelig kommunikasjon et løft. Ved å dele forskningsresultater gjennom åpne og fritt tilgjengelige publikasjoner har man muligheten til å spre forskningsresultater raskt, både til forskningskolleger og et større publikum med interesse for forskningen.

### 3. Forskningsspørsmål

”Institusjonelt arkiv – en kanal med impact?” Hovedformålet mitt har vært å undersøke om innføringen av IA har hatt en positiv effekt på gjenfinning og bruk av institusjonens forskningsproduksjon – gir IA en verdiøkning? Er IA et verktøy som utgjør en forskjell for akademiske institusjoners samlinger, et gode for synliggjøring av forskningen og institusjonen?

Med utgangspunkt i Open Access konseptet, tidligere forskning og visjonen og målene til Forskningsrådet har jeg gjennomført en studie hvor jeg har sett nærmere på publisering og tilgjengeliggjøring av vitenskaplig og faglig materiale i den norske universitets- og høgskolesektoren. I tillegg har jeg brukt tall fra Sverige for å få en sammenligning. Spesifikt har jeg sett på innføring og bruk av institusjonelt arkiv (IA) som ny formidlingskanal.

Min hypotese er at vitenskaplige artikler produsert ved norske universitet og høyskoler får økt ”citation impact” – flere siteringer, dersom artiklene er åpne og fritt tilgjengelige gjennom et IA, enn om de ikke er det.

Spørsmålene jeg stiller er:

1. I hvilken grad vil antall siteringer for artikler som er parallellpublisert i IA øke sett i forhold til artikler som kun er publisert i tradisjonelle kanaler?
2. I hvilken grad vil artikler som er parallellpublisert i IA bli sitert raskere sett i forhold til artikler som kun er publisert i tradisjonelle kanaler?

For å besvare disse spørsmålene har jeg testet ut en metode som jeg ikke kan se har vært brukt tidligere i forbindelse med IA og måling av ”impact”/anerkjennelse. Mye av arbeidet mitt har dermed vært knyttet til det metodiske, og jeg har vært like opptatt av gjennomføringen som av de faktiske analyseresultatene. IA som formidlingskanal for vitenskaplig produksjon er forholdsvis ny sett i et forskningsperspektiv, og implementering og bruk av IA er fortsatt bare i oppstartsfasen i den norske universitets- og høgskolesektoren. Tilgjengelig datagrunnlag for Norge er pr dato begrenset, slik at analyser som foretas i dag kun kan belyse utviklingstendenser. Metodikk har derfor vært viktig for meg, og jeg stiller også spørsmålet om metoden jeg bruker egner seg til å belyse de problemstillingene jeg stiller i form av spørsmålene over. Metoden og mine erfaringer blir presentert i de neste kapitlene.

#### 4. Praktisk tilnærming - metode

Bibliometri betyr «måling av litteratur», og er den metoden som er mest anvendt i studier av forskning. Det dreier seg i hovedsak om statistiske analyser av faglige siteringer og referanser, og består av at man teller 1) antall vitenskapelige artikler pr forsker, 2) hvor ofte de vitenskapelige artiklene henvises til av andre forskere og 3) hvor mange av artiklene som trykkes i de mest prestisjefulle tidsskriftene.

I denne undersøkelsen har jeg brukt en kvantitativ metode hvor jeg har konsentrert meg om et utvalg artikler og deres siteringer. Jeg har sjekket siteringshastighet ved å telle antall siteringer over en 10-årsperiode for fagfellevurderte tidsskriftsartikler, summert siteringene på ulike måter og sammenlignet tallene for artikler som er parallellpublisert i IA med tallene for artikler som kun er publisert i tradisjonelle kanaler. Jeg har i denne undersøkelsen kun vært interessert i de kvantitative resultatene og vurdert siteringshastigheten som kriterium for ”research impact” og anerkjennelse, dvs. et mål for nytteverdi, og ikke et mål for kvalitet.

De fleste som foretar undersøkelser i tilknytning til siteringer og institusjonsarkiv tar utgangspunkt i et utvalg med vitenskapelige tidsskrifter og teller siteringer pr artikkel og sjekker om disse artiklene er parallellpublisert i IA eller ikke (Lawrence 2001, Antelman 2004, Brody & Harnad 2004, Harnad et al. 2004). Jeg har valgt å snu på utvalget og har tatt utgangspunkt i det som allerede er publisert i IAer. Jeg har telt opp og beregnet siteringer for disse artiklene og sammenlignet med siteringer for artikler av samme forfatter som ikke er parallellpublisert.

Både ved å bruke den ”vanlige” metoden og den ”nye” metoden får man undersøkt ”research impact”, eller anerkjennelse, sett i lys av IA som kanal for parallellpublisering. Resultatene bør være i overensstemmelse, men det er veien til målet som er forskjellig. Den ”vanlige” metoden fokuserer, som jeg ser det, på tidsskriftene og deres forlag. I den ”nye” metoden fokuserer jeg på enkeltforskerne, deres institusjoner og produksjon. I tidsskriftene blir det publisert artikler fra mange forfattere som er ansatt ved mange ulike forskningsinstitusjoner i mange land. For å få en artikkel på trykk i de tidsskriftene som blir vurdert som prestisjetunge og viktige innen et fagfelt, og som oppfattes som en kanal som gir anerkjennelse, er forfatterne allerede silet, dvs. kun et fåtall forfattere får sine artikler publisert. Man kjenner ikke bakteppet til artiklene, dvs. man vet ikke om artikkelforfatterne og/eller de institusjonene de er ansatt ved har en positiv eller negativ holdning til OA og IA. Man risikerer å

sammenligne siteringer på artikler fra ulike forfattere og institusjoner med veldig ulik grad av allerede oppnådd anerkjennelse. Jeg vil anta at det er ulik forventning til og oppfatning av anerkjennelse knyttet til artikler fra for eksempel Harvard i USA og Handelshøgskolen eller BI i Norge.

Det som er nevnt over er faktorer som etter min mening gir støy i datagrunnlaget, og som i mindre grad får frem den faktiske effekten IA kan gi. Når jeg konsentrerer meg om den enkelte forsker og hans/hennes fagfelt innen en institusjon får jeg et mer homogent datagrunnlag. I den ”nye” metoden sammenligner jeg forfattere, faggrupper og institusjoner med seg selv, dvs. at siteringer for artikler som både er parallellpublisert i IA og artikler som ikke er det fra samme forfatter sammenlignes. Dette kan skaleres slik at man sammenligner siteringer fra flere forfattere / faggrupper innen samme institusjon, eller fra flere forfattere / faggrupper fra flere institusjoner. Det er mulig å gå videre i mitt datamateriale for å analysere hvilke tidsskrifter som er involvert og for eksempel se på forlagenes holdninger til OA, men det har jeg ikke gjort i denne undersøkelsen.

Med forfattere/forskere, fagfelt og institusjoner som fokusområde mener, og tror jeg, at resultatene ved bruk av den ”nye” metoden lettere vil gi indikasjoner på utviklingstendenser sett i lys av ”what’s in it for me”-perspektivet. Når man snakker om IA er det viktig å huske at det ikke er snakk om enten eller. Forskerne skal fortsette å publisere i andre kanaler, dvs. i tradisjonelle tidsskrifter eller i Open Access tidsskrifter, men spørsmålet er om parallellpublisering kan gi bedre synlighet, flere siteringer, og dermed økt anerkjennelse. Jeg håper min undersøkelse og metode kan være et bidrag i den sammenheng.

Siden jeg har ønsket å se på norske forhold er utgangspunktet i innsamlingen av data et sett med norske forskere fra et antall norske universitet og høyskoler og deler av deres vitenskaplige produksjon. Jeg har tatt et tilsvarende utvalg fra svenske universiteter for å få et visst sammenlignings-/kontrollgrunnlag. Artikler og navn på forskere henter jeg basert på data innhøstet til NORA fra norske IAer. Tilsvarende i Sverige fra svenske IAer som er søkbart i fellestjenesten DiVA. Siteringsdataene henter jeg fra ISI Web of Knowledge (ISI). En overordnet presentasjon av prosjektets materiale gis i kapittel 4.1.

Som tidligere nevnt er tilgjengelig datagrunnlag for Norge pr dato begrenset, slik at metodikk har vært vel så viktig som analyseresultatene. Grunnet det begrensede datagrunnlaget gir ikke

resultatene jeg kom frem til entydig ja/nei-svar, men de antyder tendenser. Ved å telle siteringer på et antall artikler for et antall forskere parallellpublisert i IA, og sammenligne det med siteringer på det samme antall artikler for de samme forskerne som ikke er parallellpublisert, samt se på utviklingen av antall siteringer før og etter innføring av IA, gir det noen indikasjoner på hvilken retning utviklingen går. Analysene og resultatene er presentert i kapittel 5.

#### 4.1 Prosjektets materiale

Datagrnnlaget for undersøkelsen og analysene mine kommer fra tre databaser som alle er tilgjengelige via Internett, slik at alt datamateriale er hentet inn elektronisk. Videre tilrettelegging av dataene for å kunne beregne, analysere og sammenligne er gjort manuelt i Microsoft Office Excel 2007. De tre basene jeg benyttet er:

- **NORA** – en nasjonal søketjeneste for åpne institusjonelle arkiv i Norge. NORA er en innhøster, dvs. at tjenesten høster/henter metadata fra IAer hos universiteter, høyskoler og det øvrige forskningsmiljøet i Norge, for så å tilby en felles søketjeneste. NORA blir i sin tur høstet av internasjonale søketjenester som for eksempel den amerikanske OAIster og den tyske Bielefeld Academic Search Engine (UiO 2009).
- SwePub (2008) – den tilsvarende nasjonale tjenesten i Sverige, men tjenesten er fortsatt under utvikling og ikke offisielt åpnet for søk enda. Prosjektperioden er satt til å være fra 2007 til 2009, og i siste planoppdatering er produksjonsrelease for søketjenesten satt til 15. November 2009. Jeg har derfor benyttet **DiVA** (Digitala Vetenskapliga Arkivet) som er en systemløsning for et IA, men som også har et felles grensesnitt for alle institusjonene som benytter løsningen. Det betyr at det er mulig å søke på tvers av institusjonsgrensene og på ulike nivåer, for eksempel på avdelings- og fakultetsnivå og ulike vitenskapsdisipliner (Uppsala 2009).
- **ISI Web of Knowledge** – en portal fra Thomson Scientific (fotnote 7) hvor flere databaser med informasjon fra ulike vitenskaplige områder og disipliner er integrert, slik at den som trenger informasjon kan utføre et felles søk for å finne alle relevante objekter uavhengig av hvilken av basene objektet opprinnelig tilhører. ISI Web of Knowledge (fotnote 9) inneholder mulighet for referansesøk, navigasjonskart for siteringer og analyseverktøy. Jeg har benyttet portalen ISI Web of Knowledge og

databasene Science Citation Index Expanded (SCI-Expanded), Social Sciences Citation Index (SSCI) og Arts & Humanities Citation Index (A&HCI) for å hente artikler og siteringsdata med utgangspunkt i de forfatternavnene jeg plukket fra NORA og DiVA (Thomson Reuters 2009).

Det finnes mange ulike dokumenttyper i NORA og DiVA, men jeg har konsentrert meg om fagfellevurderte tidsskriftsartikler. Jeg valgte dette for å få et håndterbart datagrunnlag, men også fordi dette er den dokumenttypen som gir størst uttelling, eller har størst betydning for anerkjennelse og ”research impact”. Fagfellevurderte tidsskrifter er i tillegg hoveddokumenttypen for indeksering av siteringsdata hos Thomson Scientific og basene i ISI Web of Knowledge. Siden IA som kanal først ble introdusert på 2000-tallet har jeg benyttet datamateriale for tidsperioden 1997-2008 for å få en best mulig balanse på antall dokumenter før og etter innføring av IA ved institusjonene. Siden det er tidsskriftsartikler undersøkelsen min fokuserer på, vil ordet artikkel bli brukt videre i oppgaven.

Mitt utgangspunkt har vært å undersøke forholdene i Norge og hva som har skjedd ved et utvalg norske forskningsinstitusjoner ved innføring av IA. Ved gjennomgang av NORA fant jeg at universitetene i Bergen (UiB), Oslo (UiO) og Tromsø (UiT), samt Høgskolen i Telemark (HiT) var de institusjonene som hadde flest registrerte tidsskriftsartikler, og jeg valgte å gå videre med artiklene fra disse institusjonene. Da jeg hadde samlet alle dataene fra hver av institusjonene og hentet siteringsdata fra ISI viste det seg at oppstartstidspunktene for IA ved hver av institusjonene laget stor ubalanse på artikkel- og siteringsfordeling – se Tabell 1. Spriket i oppstartstidspunktene gjorde sammenligningen mellom institusjonene og samlet på nasjonalt nivå vanskelig. For eksempel ble DUO hos UiO offisielt startet tidlig i 2002, mens TEORA hos HiT ikke startet før desember 2007. Når jeg hentet ut artikler fra perioden 1997-2008 ble det derfor en betydelig skjevfordeling på periodene før og etter innførings-tidspunktet for IA, og jeg måtte redusere utvalget i Norge til kun å gjelde universitetene i Bergen og Oslo, da de hadde oppstart nærmest hverandre i tid. Oppstartstidspunktet for disse var også det som var nærmest oppstartstidspunktet for IA ved universitetene i Uppsala og Umeå, som jeg valgte som grunnlag for sammenligning/kontroll. Metodisk er manglene i datagrunnlaget et problem i forhold til måling av statistisk representativitet, noe jeg kommer tilbake til senere i oppgaven. For hver dag som går blir det lenger siden oppstart, og i fremtidige undersøkelser vil problemet bli mindre. Flere institusjoner kan inkluderes i



analysen, fordi flere institusjoner har artikler publisert i sammenlignbare perioder før og etter innføring av IA.

Nasjonal IA	NORA				DiVA	
Navn lokalt IA	BORA	DUO	MUNIN	TEORA	Uppsala	Umeå
Institusjon	Universitetet i Bergen (UiB)	Universitetet i Oslo (UiO)	Universitetet i Tromsø (UiT)	Høgskolen i Telemark (HiT)	Universitetet i Uppsala	Universitetet i Umeå
Oppstartstidspunkt er gitt som svar på forespørsel i epost	03.11.2004 ): 2005	Oppstart i 2002 ): 2002	21.09.06 ): 2007	19.12.07 ): 2008	Okt 2001 ): 2002	2002

Tabell 1 – Oversikt over nasjonale og lokale IAer og deres offisielle oppstartstidspunkt.

For å kunne gruppere dataene og ha mulighet til å sammenligne på flere nivåer enn bare totaler for institusjoner og land valgte jeg å hente ut data fra tre ulike emne kategorier basert på emnelister gitt i NORA. Emnekategoriene som ble valgt var Samfunnsvitenskap (SAMF), Matematikk og naturvitenskap (MAT) og Medisinske fag (MED). DiVA har ikke akkumulerte data på emnenivå på samme måte, men jeg valgte uttrekk på fakultet- og avdelingsnivå som korresponderte. Kortformen for emnekategoriene, SAMF, MAT og MED blir brukt videre i analysene og presentasjonen av disse.

Uttrekk NORA&DiVA	UiB		UiO		NORGE		Uppsala		Umeå		SVERIGE		TOTALT	
	Nora	Uttrekk	Nora	Uttrekk	Nora	Uttrekk	DiVA	Uttrekk	DiVA	Uttrekk	DiVA	Uttrekk	N/D	Uttrekk
<b>SAMF</b> - ant. artikler	7	18	87	97	94	115	750	427	750	309	1500	736	1594	851
Ant. forfattere	36	9	44	35	80	44	536	25	687	29	1223	54	1303	98
<b>MAT/NAT</b> - ant. artikler	62	149	6	23	68	172	750	617	750	479	1500	1096	1568	1268
Ant. forfattere	179	36	19	19	198	55	1195	29	1336	32	2531	61	2729	116
<b>MED</b> - ant. artikler	20	28	2	6	22	34	750	338	750	377	1500	715	1522	749
Ant. forfattere	99	13	9	6	108	19	1540	26	1770	32	3310	58	3418	77
<b>TOTALT</b> - ant. artikler	89	195	95	126	184	321	2250	1382	2250	1165	4500	2547	4684	2868
Ant. forfattere	314	58	72	60	386	118	3271	80	3793	93	7064	173	7450	291
Snitt ant artikler pr forf	0,3	3,4	1,3	2,1	0,5	2,7	0,7	17,3	0,6	12,5	0,6	14,7	0,6	9,9

Tabell 2 – Oversikt over uttrekk av datamaterialet fra NORA og DiVA.

Tabell 2 viser oversikt over uttrekk av datamaterialet fra NORA og DiVA. Uttrekket fra NORA er gjort 27.01.09 og uttrekket fra DiVA er gjort 23. 03.09. Tidsforskjellen i uttrekket har ikke medført spesielle skjevheter i grunnlaget, fordi uttrekket fra Sverige måtte reduseres ytterligere for å få en håndterbar mengde artikler. Dette vil jeg komme tilbake til.

For tidsperioden 1997-2008 fant jeg 89 fagfelleverderte artikler for UIB, som fordelte seg med 7 artikler innen emne kategorien SAMF, 62 artikler innen MAT og 20 artikler innen

MED. Disse 89 artiklene er registrert med 314 forfattere/medforfattere, som gir et snitt på 0,3 artikler pr forfatter. Tallene er markert i grått i første tallrekke i kolonnen for UiB i Tabell 2. Totalt utgjorde uttrekket for de 2 norske institusjonene 184 fagfelleverderte tidsskriftsartikler innen de 3 emnekategoriene SAMF, MAT og MED. Til sammen var disse 184 artiklene registrert med 386 forfattere/medforfattere, som gir et snitt på 0,5 artikler pr forfatter. Tallene er markert i grått i første tallrekke i kolonnen for Norge i Tabell 2.

Til sammenligning er man kommet ganske mye lenger med parallellpublisering/egenarkivering i Sverige. Et tilsvarende uttrekk fra DiVA for de 2 valgte institusjonene Uppsala og Umeå var på hele 43.390 artikler. Dette ble for omfattende for min undersøkelse, så jeg reduserte utvalget og hentet 1.500 artikler pr emnekategori for hver av institusjonene, totalt 4.500 artikler. Disse var registrert med til sammen 7.064 forfattere/medforfattere, som gir et snitt på 0,6 artikler pr forfatter. Tallene er markert i grått i første tallrekke i kolonnen for Sverige i Tabell 2. Så til tross for et mye større volum er snittet på antall artikler pr forfatter tilnærmet lik, dvs. henholdsvis 0,5 for Norge og 0,6 for Sverige.

For hvert av artikkeluttrekkene, dvs. pr emnekategori pr institusjon, hentet jeg ut alle navnene på forfatterne/medforfatterne og laget forfatteroversikter hvor jeg ordnet oversiktene etter hvor mange artikler hver forfatter var registrert med. De forfatterne med flest artikler ble mitt forfatterutvalg som igjen ble brukt for videre søk mot siteringsbasen ISI. Hvor mange artikler hver forfatter er registrert med i NORA og DiVA varierte fra 1 til 70 i mitt utvalg. Det var derfor vanskelig å finne et spesifikt antall i forhold til hvor mange forfattere pr emnekategori pr institusjon jeg skulle plukke for videre søk mot ISI. Grensen på antall artikler en forfatter måtte være registrert med for å bli med i forfatterutvalget ble av den grunn ulik hos de forskjellige institusjonene, men de forfatterne som ble plukket var de med flest registreringer. For eksempel ble grensen satt på minimum 2 artikler for forfattere tilknyttet UiB, og antall forfattere her ble totalt 58, mens for Uppsala som hadde et mye større datagrunnlag satte jeg grensen til 8 artikler pr forfatter, for å få et jevnere antall forfattere pr kategori, og endte opp med totalt 80 forfattere. Totalt plukket jeg 291 forfatternavn basert på de rangerte listene pr emnekategori pr institusjon. Dette fordelt på 118 norske og 173 svenske forfattere, som ble brukte for videre søk og uttrekk av siteringsdata fra ISI. Vedlegg 8.1 viser eksempel på en ordnet fordeling av antall artikler pr forfatter.

De 291 forfatternavnene var registrert som forfattere/medforfattere for 2868 artikler av de totalt 4684 jeg i utgangspunktet hadde plukket. Det fordelte seg med 118 forfattere med 321 artikler fra NORA og Norge, og 173 forfattere og 2547 artikler fra DiVA og Sverige. Tallene er markert i grått i andre tallrekke i kolonnene for Norge, Sverige og Totalt i Tabell 2. Gjennomsnittlig antall artikler pr forfatter har økt til 9,9 (2,7 for norske forfattere og 14,7 for svenske). Årsaken til økningen er resultatet av hvilke forfattere som til slutt ble med i forfatterutvalget. Forfatterne ble plukket fra de ordnede forfatteroversiktene, og da oversiktene ble satt opp ble en artikkel som har flere forfattere gjentatt det antall ganger artikkelen har forfattere. Dersom flere av en artikkels forfattere inngikk i det antall forfattere som ble plukket til forfatterutvalget ble artikkelen også med flere ganger, nærmere bestemt det antall ganger artikkelens ulike forfattere er med i utvalget. Man ser dette tydelig for Norge hvor utgangspunktet var 184 artikler med totalt 386 forfattere. Etter utplukk ble dette 321 artikler for de 118 forfatterne i forfatterutvalget, dvs. utvalget inneholder flere forfattere som er registrert som forfatter/medforfatter til en og samme artikkel.

De 291 navnene i forfatterutvalget ble grunnlaget for søk i ISI og utplukk av de årsfordelte listene med artikler og siteringsdata pr forfatter. Dette utgjorde artikkelutvalget som ble brukt videre i alle analysene i denne oppgaven. Eksempel på oppsett av søk mot ISI er gitt i vedlegg 8.2 Forfatterutvalget - utgangspunkt for søk i ISI, og eksempel på resultatet av uttrekk fra ISI er gitt i vedlegg 8.3 Artikkelutvalget - uttrekk fra ISI.

Uttrekk ISI	UiB		UiO		NORGE		Uppsala		Umeå		SVERIGE		TOTALT	
	ISI	Match	ISI	Match	ISI	Match	ISI	Match	ISI	Match	ISI	Match	ISI	Match
<b>SAMF</b> - ant. artikler	184	0	294	20	478	20	1115	173	863	118	1978	291	2456	311
Ant. forfattere	9	0	35	16	44	16	25	21	29	26	54	47	98	63
<b>MAT/NAT</b> - ant. artikler	808	69	229	23	1037	92	2248	384	1471	235	3719	619	4756	711
Ant. forfattere	36	22	19	19	55	41	29	27	32	29	61	56	116	97
<b>MED</b> - ant. artikler	472	44	124	0	596	44	1715	183	1587	185	3302	368	3898	412
Ant. forfattere	13	11	6	0	19	11	26	26	32	31	58	57	77	68
<b>TOTALT</b> - ant. artikler	1464	113	647	43	2111	156	5078	740	3921	538	8999	1278	11110	1434
Ant. forfattere	58	33	60	35	118	68	80	74	93	86	173	160	291	228
Snitt ant artikler pr forf	25,2		10,8		17,9		63,5		42,2		52,0		38,2	
Artikler med "ny giv"/Ant forf	37	12	5	5	42	17	295	35	20	13	315	48	357	65

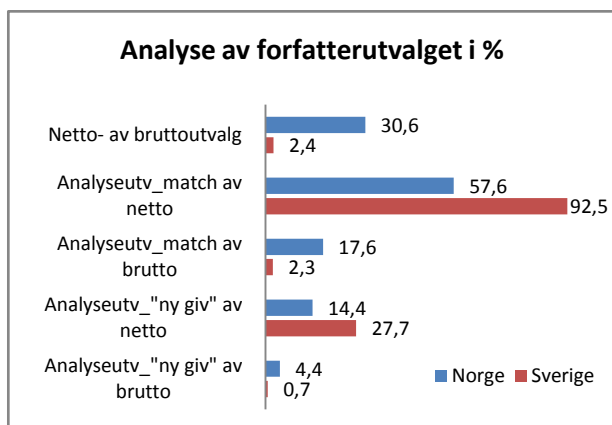
Tabell 3 - Oversikt over uttrekk av datamaterialet fra ISI.

Tabell 3 viser oversikt over uttrekk av datamaterialet fra ISI. Uttrekket er gjennomført 5.-9.2 for UiB, 16.3 for UiO og 26.-27.3 for Uppsala og Umeå.

For tidsperioden 1997-2008 fant jeg 2.111 fagfellevurderte artikler i ISI for de totalt 118 norske forfatterne, som fordelte seg med 478 artikler av 44 forfattere innen emnekategorien SAMF, 1.037 artikler av 55 forfattere innen MAT og 596 artikler av 19 forfattere innen MED. I ISI er altså snittet 17,9 artikler pr forfatter i motsetning til analyseutvalget i NORA hvor snittet er 2,7 artikler pr forfatter. Tallene er markert i grått i første tallrekke i kolonnen for Norge i Tabell 3.

For Sverige var tilsvarende uttrekk på 8.999 artikler for de 173 svenske forfatterne, som i ISI gir et snitt på 52,0 artikler pr forfatter mot DiVA hvor snittet er 14,7 artikler pr forfatter i analyseutvalget. Tallene er markert i grått i første tallrekke i kolonnen for Sverige i Tabell 3.

Når man ser nærmere på analyseutvalgene mine i Figur 1 og Figur 2 nedenfor, og gjør en vurdering av disse opp mot både netto- og bruttoutvalgene (forklart i gjennomgangen av Figur 1 og Figur 2), så utgjør analyseutvalgene en så liten andel av netto- og bruttoutvalgene at det er stor sannsynlighet for at de ikke er statistisk representative.



**Figur 1 – analyse av forfatterutvalget.**

#### Gjennomgang av Figur 1.

Bruttoutvalget i denne analysen er antall forfattere totalt i mitt utvalg hentet fra NORA/ DiVA innen emnekategoriene MAT, MED og SAMF.

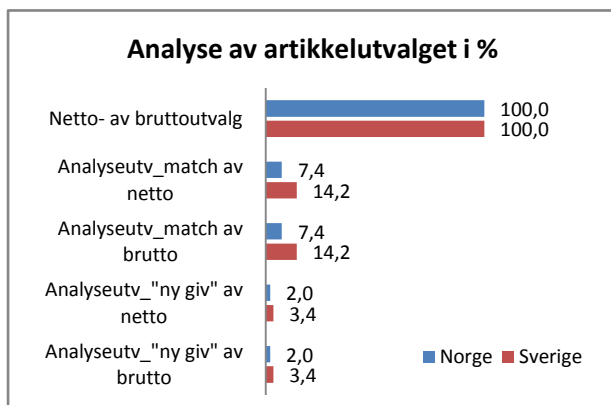
(Norge:386 Sverige:7064)

Nettoutvalget er antall forfattere plukket i NORA/DiVA for videre søk i ISI.

(Norge:118 Sverige:173)

Analyseutvalget\_match er antall forfattere med artikler som har match i NORA/DiVA og ISI, dvs. at artiklene finnes både i NORA/DiVA og ISI. (Norge:68 Sverige:160)

Analyseutvalget\_'ny giv' er antall forfattere med artikler som har match i NORA/DiVA og ISI, og hvor artiklene er publisert minimum 3 år før innføring av IA. (Norge:17 Sverige:48)



Figur 2 – analyse av artikkelutvalget.

## Gjennomgang av Figur 2.

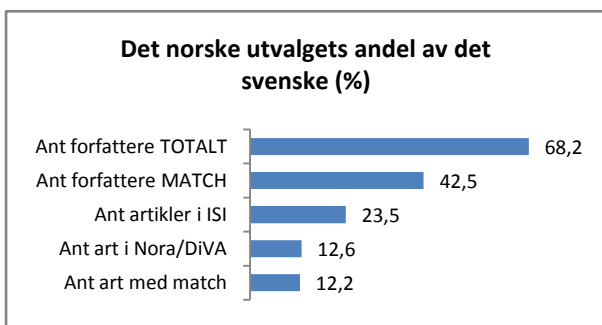
Bruttoutvalget i denne analysen er antall artikler totalt i mitt utvalg hentet fra ISI for valgte forfattere innen emnekategoriene MAT, MED og SAMF. (Norge:2111 Sverige:8999)

Nettoutvalget er det samme som bruttoutvalget. Alle artiklene som ble hentet fra ISI for valgte forfattere ble brukt i videre analyser. (Norge:2111 Sverige:8999)

Analyseutvalget\_match er antall artikler i ISI som har match i NORA/DiVA, dvs. at artiklene finnes både i NORA/DiVA og ISI. (Norge:156 Sverige:1278)

Analyseutvalget\_'ny giv' er antall artikler i ISI som har match i NORA/DiVA, og hvor artiklene er publisert minimum 3 år før innføring av IA. (Norge:42 Sverige:315)

Det er ingen systematisk forskjell mellom bruttoutvalg og nettoutvalg, og frafallet er tilfeldig. Det begrunner jeg med at jeg tar utgangspunkt i alle artiklene for de valgte emne kategorier jeg fant i NORA, dvs. 100 %, mens jeg for Sverige måtte ta et plukk på de 1500 første artiklene pr kategori for å håndtere omfanget i en videre manuell behandling. For Sverige benytter jeg derfor kun 10,4 % av totalt antall artikler. Basert på dette utgangspunktet ble alle forfatterne ordnet etter antall artikler de er registrert med i NORA/DiVA, uavhengig om forfatteren var hovedforfatter eller medforfatter. Videre ble de forfatterne med flest antall artikler plukket ut for videre søk i ISI Web of Knowledge. Også antallet forfattere ble tilfeldig basert på å få et visst antall forfattere pr kategori pr institusjon, men jeg ser i ettertid at det burde vært en prosent-andel av forfatterne og ikke basert på mengden artikler. Utplukket av forfatterne er kun basert på eksistensen av artikler innen valgte emne kategorier og institusjoner. Jeg har ikke benyttet demografiske variable som for eksempel alder, kjønn, utdanning, ansettelsestid eller lignende. Jeg har heller ikke vurdert vektning for å veie opp for mangler i utvalget, men valgt å basere beregningene i analysene på snittberegninger og prosentvise fordelinger (Ringdal 2007).



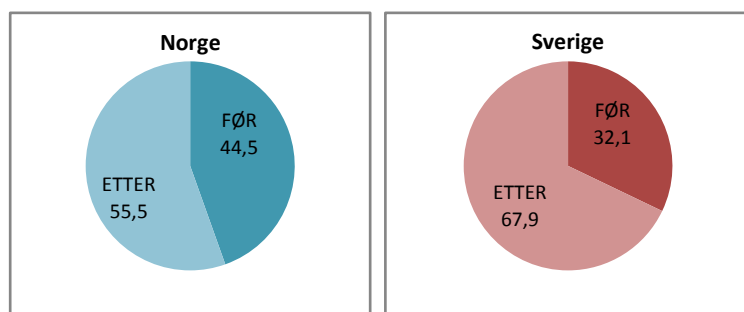
**Figur 3 - forskjeller i datagrunnlaget. Her vist med hva det norske utvalget utgjør av det svenske utvalget, dvs. andel i %.**

Forskjellen mellom det norske og svenske utvalget er stort – illustrert i Figur 3. Da jeg gikk inn i det mulige datagrunnlaget for Norge med utgangspunkt i NORA ble det klart at fenomenet IA er nytt, og at antall artikler registrert i de enkelte arkivene fortsatt er få. Selv om alle de 7 norske universitetene og 24 av 42 høgskolene har

innført IA er mange av disse nylig startet, og antall artikler er deretter. Sverige har kommet adskillig lengre når det gjelder antall artikler for de institusjonene som benytter DiVA, men jeg kjenner ikke tallet for spredningen av IA blant Sveriges totale antall universitet og høgskoler.

I det følgende beskriver jeg dataene i NORA/DiVA og ISI mer i detalj, viser sammenhenger og forklarer relasjonene. En oversikt med grunnlaget for grafene er gjengitt i vedlegg 8.4 Sammenstilling av uttrekk fra NORA/DiVA og ISI.

#### Uttrekket fra NORA/DiVA.



**Figur 4 – antall artikler i uttrekket fra NORA og DiVA fordelt på tidsperiodene FØR og ETTER innføring av IA.**

44,5 % av artiklene jeg fant i NORA for de 118 norske forskere jeg valgte var publisert før innføringstidspunktet av IA, mens 55,5 % var publisert etter – se Figur 4. Tilsvarende var 32,1 % av artiklene i DiVA for ”mine” 173 svenske forskere publisert før innførings-tidspunktet av IA, mens 67,9 % var publisert etter.

Det at andelen av artikler er større både i NORA og DiVA for artikler publisert etter innføring i IA er for så vidt ikke overraskende, da det er naturlig å starte registrering på tidspunktet for oppstart. Derfor er det mer interessant å se at så mange artikler som henholdsvis 44,5 % og 32,1 % er fra perioden før oppstart av IA. Det betyr at institusjonene har lagt seg på en linje

med å etterregistrere flere av de eldre artiklene for å få fulltekstdokumentene tilgjengelig, samtidig som samlingen av fulltekstdokumenter blir mer fullstendig. Jeg kan bare spekulere i hvilke artikler dette gjelder, fordi jeg ikke har sjekket dette nærmere. For min undersøkelse ville det være av interesse å vite om dette var høyt anerkjente artikler plukket ut spesielt for etterregistrering, eller om det var artikler som forskerne selv har valgt å etterregistrere av egeninteresse.

### Uttrekket fra ISI.

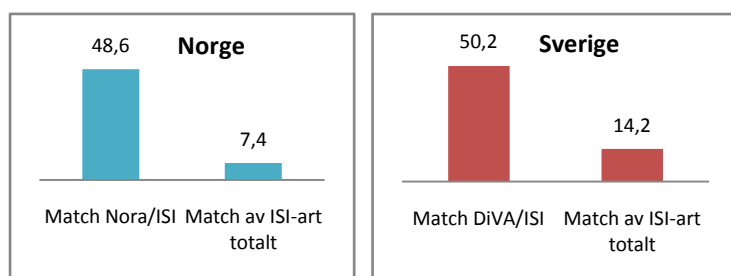


**Figur 5 - antall artikler i uttrekket fra ISI fordelt på tidsperiodene FØR og ETTER innføring av IA.**

53,5 % av artiklene jeg fant i ISI for ”mine” 118 norske forskere var publisert før innføring av IA, mens 46,5 % var publisert etter – se Figur 5. Tilsvarende for ”mine” 173 svenske forskere var 40,0 % publisert før innføring av IA, mens hele 60,0 % var publisert etter.

Her er det interessant at forskjellen på antall artikler før og etter årstallet for innføring av IA er så stor mellom Norge og Sverige. For Norge er kun 46,5 % av artiklene publisert etter, mens det for Sverige er hele 60,0 %. Uten at jeg har gått inn i datagrunnlaget kan jeg igjen bare spekulere, men en av årsakene kan være at svenskene har dreid publiseringen de senere årene til å publisere mer i internasjonale tidsskrifter enn hva norske forskere gjør. Det er svært få ikke-engelskspråklige tidsskrifter som blir indeksert i ISI, og dermed få av de nordiske tidsskriftene som er hovedpubliseringskilden for norske forskere.

### Sammenhengen mellom uttrekkene fra NORA/DiVA og ISI.



**Figur 6 – antall artikler som er felles for NORA/DiVA og ISI, dvs. det er en match, og hvor stor andel disse utgjorde av artiklene totalt i artikkelutvalget fra ISI.**

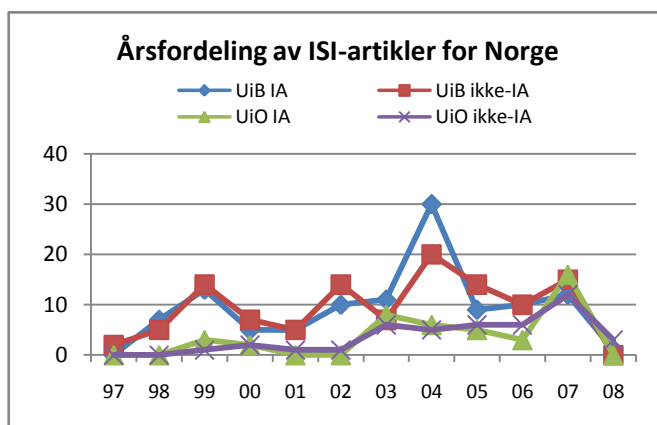
48,6 % av artiklene i NORA fant jeg også i ISI, dvs. det var match. Likevel utgjorde disse artiklene kun 7,4 % av artiklene jeg fant totalt for ”mine” norske forfattere i ISI – se Figur 6. Tilsvarende var

det 50,2 % av artiklene i DiVA som hadde match i ISI, og disse utgjorde 14,2 % av artiklene jeg fant for ”mine” svenske forfattere.

Forskjellen på andel artikler med match mellom NORA og ISI på den ene siden og DiVA og ISI på den andre siden er ikke vesentlig når man leser de prosentvise verdiene ut fra grafen her, henholdsvis 48,6 % for Norge og 50,2 % for Sverige. Likevel utgjør det norske artikkelutvalget med match bare 12,2 % av det svenske utvalget – illustrert i Figur 3 som viser forskjeller i datagrunnlaget. Det betyr at Sverige har over åtte ganger så stor andel artikler med match mellom DiVA og ISI enn det er mellom NORA og ISI. Det har klart en sammenheng med at artikkeluttrekk fra NORA kun utgjør 12,6 % av artikkeluttrekket fra DiVA, og at artikkeluttrekket for Norge i ISI kun utgjør 23,5 % av artikkeluttrekket i ISI for Sverige, men hvor forskerne publiserer er antageligvis vel så viktig.

### Spesielt om artikkeluttrekket fra ISI.

Alle artiklene hentet fra ISI er publisert i tidsintervallet 1997-2008, som også er måleperioden for analysene. Det betyr at det i utvalget finnes artikler publisert allerede i 1997, som har siteringsdata for hele tidsperioden, og det finnes artikler av nyere dato, publisert helt i slutten av tidsintervallet, som har siteringsdata kun for en kort periode. Alle siteringer er summert og snittfordelt pr forfatter knyttet opp til emnekategorier, institusjoner og land. I tillegg er de gruppert i ”IA artikler”, dvs. artikler som både finnes i ISI og NORA/DiVA, og ”ikke-IA artikler”, dvs. de artiklene jeg har plukket fra ISI, som kun finnes i ISI og som ikke er parallellpublisert. Begrepene IA artikler og ikke-IA artikler blir brukt videre i oppgaven.

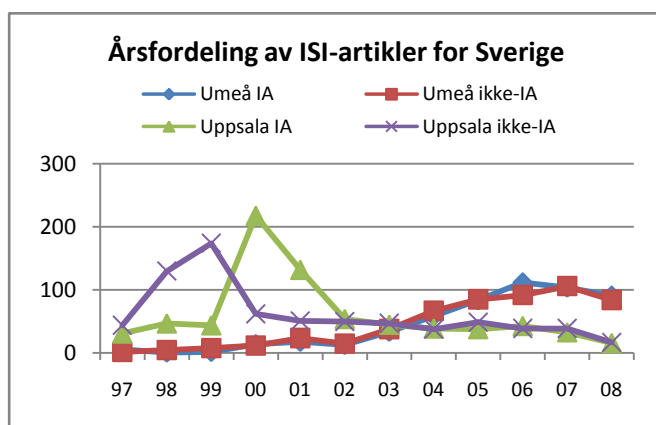


Figur 7 - årsfordeling av ISI-artikler for Norge basert på publiseringsår.

For Norges vedkommende er fordelingen av artikler i ISI, gruppert som IA og ikke-IA artikler, jevn over alle årene i tidsintervallet for undersøkelsen. Grafen i Figur 7 viser at publiseringsårene stort sett er overlappende ved at linjene følger, og nesten dekker hverandre. Utvalget for Norge med artikler som har match mellom ISI og NORA er som tidligere



nevnt lite (156 artikler), med flest artikler indekstert i 2004 (36 IA artikler og 25 ikke-IA artikler). De fleste av disse artiklene er fra emne kategorien MAT hos UiB.



Figur 8 - årsfordeling av ISI-artikler for Sverige basert på publiseringsår.

Andelen av artikler i utvalget med match mellom ISI og DiVA er høyere for Sverige enn for Norge (henholdsvis 1278 mot 156 artikler). Kurvene i Figur 8 viser at det er en parallellforskyvning hos Sverige med et høyere antall ikke-IA artikler fra perioden 1998-1999 (135 og 182 artikler) og et tilsvarende høyt antall IA artikler fra 2000-2001 (232 og 150 artikler). Disse toppene for Sverige gjelder for det meste artikler fra emne kategorien MAT hos Uppsala.

## 4.2 Statistisk representativitet

I forrige avsnitt beskrev jeg analyseutvalgene og svakhetene ved disse. Utvalgene utgjør så liten andel av netto- og bruttoutvalgene at det er stor sannsynlighet for at de ikke er statistisk representative. Tilgjengelig datagrunnlag for Norge er pr dato meget begrenset, både fordi IA som formidlingskanal er ny, og fordi rutiner knyttet til parallellpublisering fortsatt er under utarbeidelse og i innkjøringsfasen hos de fleste av institusjonene som har innført IA. Til tross for det har jeg valgt å gjennomføre flere analyser og presenterer de her i oppgaven. Det har vært vesentlig for meg at valg av metode har vært like viktig som de resultatene jeg fikk på de enkelte analysene, og jeg har vært klar på at resultatene ikke kan være eksakte fakta, men antydning av tendenser.

I tillegg til at datautvalget for IA i Norge er begrenset, har jeg redusert utvalget ytterligere i forhold til den tiden jeg har hatt til disposisjon for gjennomføring av undersøkelsen. Dataene, slik jeg samlet og organisert dem med tanke på gjennomføring av analysene, lot seg ikke gjenbruke i signifikanstester på en enkel måte. Blant annet har jeg summert og aggregert totalt antall siteringer og snitt antall siteringer i forhold til begrepene ”før IA” og ”etter IA” (detaljert beskrivelse gis i kapittel 5.15.1.1 Endringer i antall siteringer før og etter tidspunktet for

innføring av IA), og ikke for hvert enkelt år innenfor måleperioden. Med bakgrunn i dette har jeg valgt å ikke gjennomføre en signifikanstest på analyseutvalgene brukt i undersøkelsen. Isteden beskrives i det følgende en anbefaling og mulig fremgangsmåte for gjennomføring av en slik test.

- **Anbefaling av teknikk for signifikanssjekk**

Ut fra betraktning av dataenes beskaffenhet ville jeg ha valgt en ikke-parametrisk testteknikk for å sjekke dataenes signifikans. En parameter blir i statistikken definert som et mål som beskriver populasjonen. Skal man benytte en parametrisk teknikk forutsettes det at populasjonen er normalfordelt<sup>10</sup> og målingene må være tilknyttet den fullstendige informasjonen som ligger i kontinuerlige variable<sup>11</sup>. En ikke-parametrisk teknikk stiller ikke så strenge krav til fordeling og målenivå, men bygger heller på ranger eller frekvenstabeller. Ikke-parametriske tester er likevel mindre effektive enn de tilsvarende parametriske testene, fordi de bygger på færre og enklere forutsetninger, men som motsatt er de ofte lettere å forstå og mindre følsomme for ekstreme observasjoner (Ringdal 2007).

- **Anbefaling av testmetode for signifikanssjekk**

Det finnes flere ikke-parametriske testmetoder som kunne ha vært anvendt for gjennomføring av en signifikanssjekk i undersøkelser som denne, men jeg vil anbefale Friedman-testen på grunn av sin enkle og pedagogiske oppbygging.

Milton Friedman (1912-2006), amerikansk professor i økonomi med blant annet 30 år ved University of Chicago, vant i 1976 Alfred Nobels økonomipris for sine bidrag på områder som forbruksanalyse, valutahistorie og valutateori. Han var mest kjent for sine teorier innen økonomi, men har i tillegg utviklet en testteknikk som baserer seg på ikke-parametrisk statistikk – den såkalte Friedman-testen.

Friedman-testen er en ikke-parametrisk test som sjekker sannsynligheten for at forskjellene i utvalget er tilfeldig ved å sammenligne måltallene for hvert subjekt eller behandling. For å eksemplifisere ville måltallene i denne undersøkelsen kunne vært snitt

---

<sup>10</sup> **Normalfordeling** – også kalt Gausskurve, er en viktig fordeling i statistikk og sannsynlighetsteori. En normalfordelt variabel antar ofte verdien som ligger nær middelveidien, og sjelden verdien som har stor avvikelse. Normalfordelingen er symmetrisk og ser ut som en klokke.

<sup>11</sup> **Kontinuerlige variabler** – analyseenheter beskrives ved hjelp av variabler. Hvis enhetene er mennesker er kjønn, utdanning og yrke eksempler på variabler. Kontinuerlige variabler betyr at variablene kan anta et uendelig antall mulige verdier.

antall siteringer hvert av årene i måleperioden, dvs. fra 1998 til og med 2008. Subjektene vil kunne være institusjonene som er med i undersøkelsen, dvs. UiB, UiO og universitetene i Uppsala og Umeå.

- **En mulig gjennomføring av Friedman-testen**

I Friedman-testen blir de originale måltallene for hvert subjekt byttet ut med en ranking som testes mot en nullhypotese, som oftest formulert som det motsatte av det man tror<sup>12</sup>. Nullhypotesen i denne undersøkelsen kunne ha vært at en økning i antall siteringer etter innføring av IA var tilfeldig. Rankingen for hvert av subjektene summeres og det beregnes gjennomsnitt og median pr år. Disse brukes så til å beregne en statistisk testobservator ( $\chi^2$ ), som blir grunnlaget for å bestemme om nullhypotesen skal beholdes eller forkastes.

Ved en eventuell gjennomføring av signifikanssjekk basert på Friedman-testen vil jeg anbefale lenkesamlingen til Vassar College i USA. De har lagt ut lenker til en åpen og fritt tilgjengelig statistisk lærebok hvor Friedman-testen og andre testteknikker beskrives i detalj. Læreboka følger nettstedet ”VassarStats” som er et nettsted hvor teknikkene også kan testes ved å legge inn eget datamateriale (Lowry 2009a, Lowry 2009b).

## 5. Presentasjon av resultater - analyse og diskusjon

Hovedformålet med min undersøkelse har vært å undersøke om innføringen av IA har hatt en positiv effekt på gjenfinning og bruk av institusjonens forskningsproduksjon – om IA gir en verdiøkning? Min hypotese er at vitenskaplige artikler produsert ved norske universitet og høyskoler får økt ”citation impact” – flere siteringer, dersom artiklene er åpent og fritt tilgjengelige gjennom IA, enn om de ikke er det.

For å teste ut og måle min hypotese har jeg stilt to spørsmål:

1. I hvilken grad vil antall siteringer for artikler som er parallellpublisert i IA øke sett i forhold til artikler som kun er publisert i tradisjonelle kanaler?
2. I hvilken grad vil artikler som er parallellpublisert i IA bli sitert raskere sett i forhold til artikler som kun er publisert i tradisjonelle kanaler?

---

<sup>12</sup> **Poppers falsifiseringsprisnipp** – det er umulig å påvise at en hypotese er sann, men vi kan vise at den er gal. Ved å forkaste nullhypotesen styrkes troen på den alternative hypotesen (Ringdal 2007, s. 239).

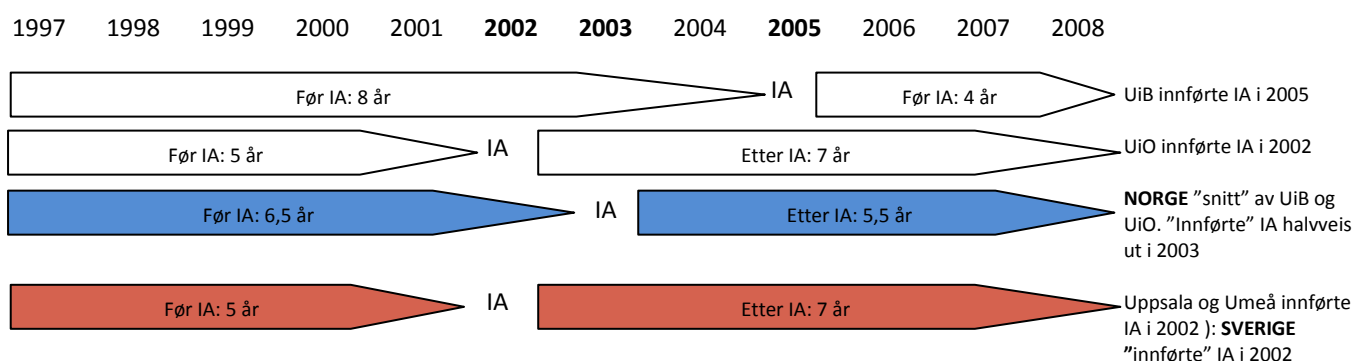
Resultatene jeg kom frem til gir som nevnt ikke entydig ja/nei-svar, og resultatene er ikke eksakte fakta, men antyder tendenser. Analysene og resultatene er presentert fortløpende i de neste avsnittene med utgangspunkt i spørsmålene over.

Siden det har vært like viktig for meg å teste en ny metode, som å finne analyseresultater, stiller jeg også spørsmålet om metoden egner seg til å belyse de problemstillingene jeg tar opp. Til slutt i dette kapitlet har jeg beskrevet mine erfaringer med metoden, og jeg kommer med anbefalinger til eventuell fremtidig bruk.

## 5.1 Endringer i antall siteringer

For å svare på det første spørsmålet; om i hvilken grad antall siteringer for artikler som er parallellpublisert i IA vil øke sett i forhold til artikler som kun er publisert i tradisjonelle kanaler, har jeg foretatt to analyser. Den første så på eventuelle endringer i antall siteringer før og etter tidspunktet for innføring av IA. Den andre analysen så på eventuelle forskjeller i antall siteringer for artikler parallellpublisert i IA sammenlignet med de som ikke er det. Resultatene er presentert og diskutert for hver av analysene i de neste underkapitlene.

Før jeg presenterer resultatene vil jeg definere begrepene ”før IA” og ”etter IA” som er sentrale i alle analysene. I de neste avsnittene og i tilknytning til forklaringen til Figur 10 i kapittel 5.1.1 er de to begrepene understreket for å fremheve uttrykkene som begrep. Senere i oppgaven blir uttrykkene brukt og skrevet i løpende tekst.



Figur 9 - visualisering av begrepene "før IA" og "etter IA".

”Før IA” og ”etter IA” angir de to tidsperiodene som måles og sammenlignes. Undersøkelsens totale tidsintervall eller måleperiode er mellom 1997 og 2008. Dette intervallet er delt i to

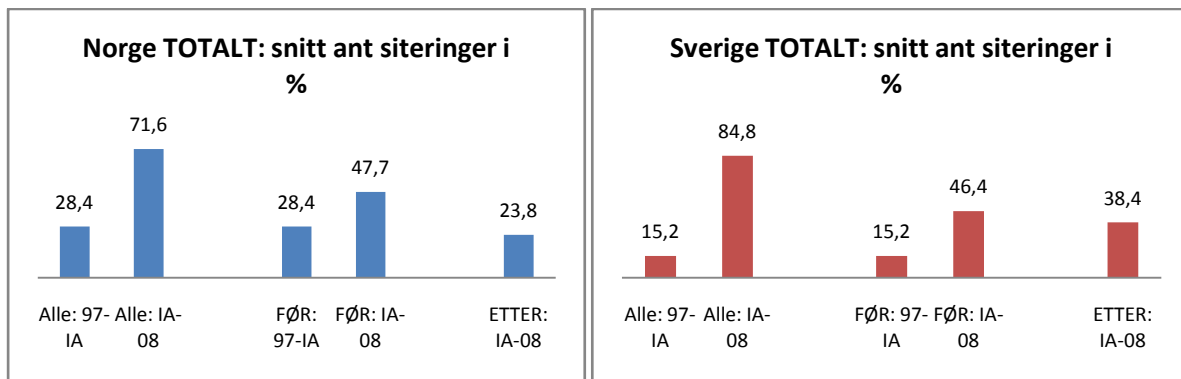
med en periode før innføring av IA og en etter innføring av IA. Jeg opererer med de relative begrepene før og etter, fordi innføringstidspunktet av IA hos de 4 institusjonene jeg har med i undersøkelsen er forskjellig. DUO ble innført hos UiO i 2002 og BORA hos UiB i 2005. Snitt i utregningene blir da midtveis i 2003, dvs. før utgjør 6,5 år for perioden 1997-2003, mens etter utgjør 5,5 år for perioden 2003-2008. For Sverige er det motsatt, dvs. færre år i perioden før enn etter. Både Uppsala og Umeå innførte DiVA i 2002, og før utgjør da 5 år for perioden 1997-2001, mens etter utgjør 7 år for perioden 2002-2008.

Perioden jeg teller siteringer for er altså litt lengre for før enn etter når det gjelder Norge. Antall artikler i ISI fordeler seg med henholdsvis 53,5 % for perioden før og 46,5 % for perioden etter – vist i Figur 5. Når det gjelder Sverige er det motsatt, dvs. perioden jeg teller siteringer for er litt kortere for før enn etter, og antall artikler fordeler seg med henholdsvis 40,0 % for før og 60,0 % for etter. Konsekvensen av dette er at analyseutvalgene mine er noe skjevt i forhold til de målingene jeg foretar med ”før IA” og ”etter IA”. Jeg vurderer forskjellene til ikke å være så vesentlige at tallmaterialet ikke kan la seg sammenligne. For å kompensere for skjevhetene foretar jeg snitt- og prosentberegninger.

### **5.1.1 Endringer i antall siteringer før og etter tidspunktet for innføring av IA**

I denne analysen tok jeg utgangspunkt i alle artiklene som jeg hentet fra ISI (totalt 11.110 artikler for både Norge og Sverige), og beregnet totalt antall siteringer og snitt pr forsker summert for artikler publisert i periodene før og etter innføring av IA. Disse beregningene er uavhengige om artiklene er parallellpublisert eller ikke. Ved å foreta en slik undersøkelse kan jeg se om det er en utvikling på antall siteringer kun basert på det faktum at IA er innført.

Jeg beregnet totaler og snitt i antall og prosent pr forsker, og summerte så dette på nivåene emnekategori, institusjon og land. De akkumulerte tallene gjorde det mulig å sammenligne og se på eventuell utvikling for grupper av forskere innen emnekategoriene og institusjonene i Norge opp mot hva som skjer i Sverige.



Figur 10 - snitt antall siteringer for artikler fra ISI fordelt på tidsperioden før og etter innføring av IA.

I de to grafene i Figur 10 viser den første gruppen søyler, **alle**, en prosentvis fordeling av snitt antall siteringer for alle artiklene funnet i ISI publisert i perioden 1997-2008 for de valgte forfatterne. Hver av de to søylene viser snitt antall siteringer før og etter innføring av IA uavhengig av når artiklene er publisert. Det er kun skille på om artikkelen er sitert før (97-IA) eller etter (IA-08) tidspunktet for innføring av IA.

De to neste søylegruppene, **før** og **etter**, er en detaljering av den første søylegruppen **alle**. Disse to søylegruppene viser snitt antall siteringer basert på artiklenes publiseringstidspunkt. Gruppen **før** viser fordeling av snitt antall siteringer for alle artiklene funnet i ISI publisert før innføring av IA. Søylen FØR:97-IA viser snitt antall siteringer gitt i perioden før innføring av IA for artikler publisert før innføring av IA. Søylen FØR:IA-08 viser snitt antall siteringer gitt i perioden etter innføring av IA for de samme artiklene, dvs. artiklene publisert før innføring av IA. Tilsvarende viser søylen ETTER:IA-08 snitt antall siteringer gitt etter innføring av IA knyttet til artiklene publisert etter innføring av IA. I disse grafene er altså snitt antall siteringer før og etter innføring av IA basert på alle artikler i ISI - uavhengig om artiklene er parallellpublisert eller ikke.

Basert på tallverdiene på søylene i Figur 10 kan det se ut som om antall siteringer øker kraftig for nye artikler. Tallene viser en klar økende tendens, men det kreves en forklaring.

- Ser man kun på søylen **alle** er økningen på antall siteringer for Norges vedkommende på hele 151,5 % for perioden etter i forhold til før innføring av IA, fra 28,4 til 71,6 %. Når man vet at antall artikler i de samme periodene er på henholdsvis 53,5 og 46,5 % for periodene før og etter, og man tar i betraktning at siteringer normalt starter et par år etter publisering, så skal man her forvente flere siteringer i perioden etter. Økningen

av siteringer er likevel vesentlig, og kan ikke forklares alene med fordelingen av artikler.

Heller ikke for Sverige kan økningen av antall artikler være hele forklaringen for den sterke økningen i snitt antall siteringer for perioden før i forhold til perioden etter innføring av IA, som er på hele 456,3 %. Fordelingen på antall artikler i utvalget er på henholdsvis 40,0 og 60,0 % for periodene før og etter, dvs. en økning i antall artikler på 50,3 %. I Sverige innførte man IA i 2002, slik at siteringene for de ”nye” artiklene, dvs. de som er publisert etter innføring av IA, sannsynligvis er tatt opp i økningen i antall siteringer i perioden. Likevel er en økning for antall siteringer på 456,3 % vesentlig i forhold til en økning på 50,3 % i antall artikler.

- Ser man nærmere på søylene for **før** viser også disse en økning i antall siteringer – her på 67,8 % for Norges vedkommende, fra 28,4 til 47,7 %. Søylene viser snitt antall siteringer i prosent for periodene før og etter IA, men for de samme artiklene, dvs. for artiklene som er publisert før innføring av IA. Her kan en forklaring på økningen være at antall siteringer øker med årene fordi artiklene har vært lengre tilgjengelig, og flere dermed har hatt mulighet til å lese og sitere. Samtidig må det presiseres at perioden som det telles siteringer for er på 6,5 år for perioden før og på 5,5 år for perioden etter. Tallet for antall siteringer bør derfor relativt sett ha vært enda høyere for perioden etter om dette var vektet. Men likevel, til tross for at siteringsperiodene ikke er vektet, er snitt antall siteringer høyere for perioden etter, noe som gir en indikasjon på at tendensen er klart økende.

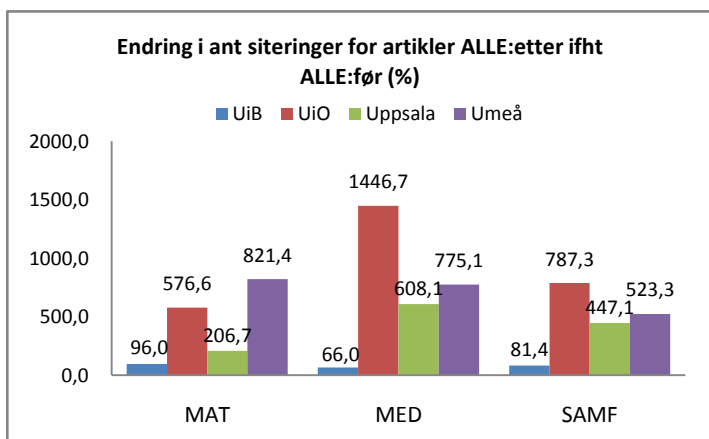
Den samme tendensen ser man for Sverige. Her er økningen på hele 204,7 % fra et snitt på 15,2 % for perioden før til 46,4 % for perioden etter. For Sveriges vedkommende kan noe av økningen forklares med at antall artikler er høyere for perioden etter innføring av IA enn før. Periodene det telles siteringer for er også lengre for perioden etter enn perioden før, henholdsvis på 7 og 5 år. Til tross for at økningen i antall siteringer kunne blitt noe redusert dersom man hadde tatt hensyn til fordeling av antall artikler og siteringsperiode gjennom vekting, så mener jeg at jeg har belegg for å si at økningen av antall siteringer på 204,7 % er vesentlig.

- Hvis man tar med den siste søylen, i gruppen **etter**, så viser den snitt antall siteringer for artikler publisert etter innføring av IA. Disse artiklene har da følgelig ikke siteringer i perioden før innføring. For Norge viser tallet snitt antall siteringer i prosent for de første 5,5 årene etter publisering for de artiklene snittet gjelder for. For å gi en viss indikasjon kan dette tallet sammenlignes med søylen FØR:97-IA, som viser prosentvis snitt siteringer for de første 6,5 årene etter publisering for de artiklene dette snittet gjelder for. For Norge ser vi her en nedgang på 16,3 % fra et snitt på 28,4 % til et snitt på 23,8 %. Nedgangen kan skyldes at både antall artikler i ISI er høyere og tidsperioden siteringene gjelder for er lengre for perioden før innføring av IA enn etter, henholdsvis 53,5 % og 6,5 år for perioden før og 46,5 % og 5,5 år for perioden etter. Ser man samtidig på selve snitt-tallet, dvs. snitt antall siteringer pr artikkel pr forfatter, for de samme periodene, er snittet på 8,4 siteringer pr artikkel for artikler publisert før innføring, mens snittet er 8,1 for artikler publisert etter. Snittet pr artikkel basert på det antall forfattere jeg har i utvalget er altså forholdsvis likt for artikler publisert før og etter innføring av IA, men det viser nedgang.

Ser man på Sverige er tendensen helt annerledes. Her har det prosentvise snittet økt med 152,2 % fra et snitt på 15,2 % til et snitt på 38,4 %. Igjen kan deler av det forklares med at antall artikler er høyere for perioden etter innføring av IA enn før, og periodene det telles siteringer for er på henholdsvis 5 år og 7 år for perioden før og etter, men tar man her frem snitt-tallet, dvs. antall siteringer pr artikkel pr forfatter, er snittet 6,1 for artikler publisert før innføring, mens snittet er økt til 10,2 for artikler publisert etter. Forskjellen er helt klart større, og viser en klar positiv tendens.

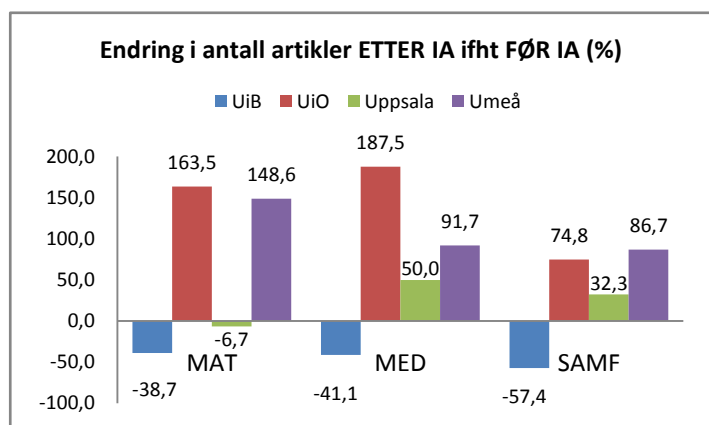
Ovenfor analysert jeg tallene på landnivå. Ser man nærmere på detaljene i Figur 11 viser også disse tallene en økning i antall siteringer innen alle emnekategoriene og alle institusjonene når man ser alle artiklene i ISI under ett, dvs. uavhengig om de er parallellpublisert eller ikke. Det som blir tydeligere er den ujevne økningen både innenfor hver emnekategori, men også innenfor hver institusjon. I emne kategorien MED har UiB lavest økning med 66,0 %, mens UiO har en økning på hele 1446,7 %. UiB har minst økning på tvers av kategoriene sammenlignet med de andre institusjonene, mens UiO og Umeå har størst økning. Økningen på hele 1446,7 % i MED for UiO kan tildels forklares med at antall artikler ble tredoblet i perioden, men når antall siteringer samtidig er 15 ganger så høy, gir det en sterk indikasjon på at noe skjer etter tidspunktet for innføring av IA.





**Figur 11 – endring i antall siteringer for artikler i ISI. Viser at antall siteringer for perioden ALLE:IA-08 (etter) er høyere enn perioden ALLE:97-IA (før).**

5 år for UiO, Uppsala og Umeå, mens den er 8 år for UiB. Perioden etter (IA-08) er 7 år for UiO, Uppsala og Umeå, mens den er kun 4 år for UiB, som nok forklarer at økningen for UiB er lavere enn for de andre institusjonene.



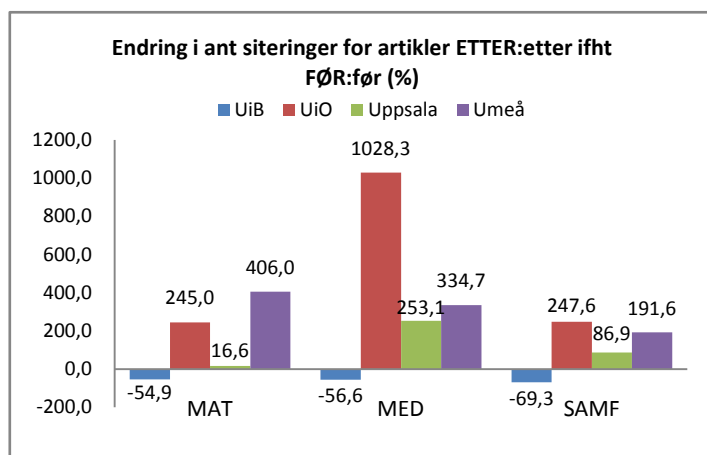
**Figur 12 - endring i antall artikler i ISI. Viser at antall artikler ETTER IA i all hovedsak er høyere enn antall artikler FØR IA.**

publisert etter IA ha kommet med i perioden det telles siteringer for, slik at en økning av antall siteringer er å forvente. Økningen i antall siteringer er likevel så mye større enn økningen i antall artikler at man ikke kan se bort fra det. For UiB blir dette enda tydeligere, fordi siteringsperioden for perioden før er på hele 8 år og kun 4 år for perioden etter (BORA ble innført i 2005). Antall artikler publisert før innføring av IA er større, og man kan forvente at siteringer for de som er publisert sent i perioden før IA først kommer frem i tallmaterialet for siteringer etter innføring av IA. Dette med årsfordeling av artiklene i ISI skal jeg komme tilbake til. Alle ting tatt i betraktning, og til tross for at økningen hos UiB er den laveste blant institusjonene i undersøkelsen, har også UiB en vesentlig økning i antall siteringer på

Det er riktig og forventet at det er en økning i antall siteringer for perioden etter innføring av IA sammenlignet med perioden før, siden perioden det telles siteringer for i all hovedsak er lengre for perioden etter innføring av IA enn perioden før. Tallene viser likevel at økningen i antall siteringer er vesentlig sett i forhold til forskjellen på antall år det beregnes siteringer for. Perioden før (97-IA) er

Forskjellen på antall artikler i ISI for den enkelte institusjon fordelt på emnekategori er vist i Figur 12. Grafen viser at alle institusjonene med unntak av UiB har flere artikler tilknyttet perioden etter IA enn perioden før IA. Siden innføringen av DUO hos UiO og DiVA hos Uppsala og Umeå var i 2002 vil siteringene for de fleste av artiklene

henholdsvis 96,0 % i emne kategorien MAT, 66,0 % i MED og 81,4 % i SAMF – vist i Figur 11.



Figur 13 – endring i antall siteringer for artikler i ISI. Viser at antall siteringer for perioden ETTER:IA-08 i all hovedsak er høyere enn perioden FØR:97-IA.

Selv om Norge totalt sett har en nedgang på 16,3 % i antall siteringer for artikler publisert etter innføring av IA i forhold til artikler publisert før, vist i Figur 10 med 28,4 % for FØR:97-IA til 23,8 % for ETTER:IA-08, så bekreftes den generelle økningen når en sammenligner siteringer for de første årene etter publisering både for artikler publisert før og etter

innføring av IA – se Figur 13. Reduksjonen for Norge skyldes i all hovedsak fall i antall siteringer for UiB, som igjen i all hovedsak skyldes antall år det telles siteringer for i de to periodene og artikkelfordelingen. Siteringsperioden for UiB er på 8 år for perioden før med 927 artikler og 4 år for perioden etter med 537 artikler. Ellers er det en økning for alle emne kategorier og alle institusjoner. Igjen kan man argumentere med at UiO, Uppsala og Umeå hadde oppstart av IA i 2002, slik at antall år det beregnes siteringer for er flere etter innføring enn før, samt at antall artikler er flere i perioden etter enn før, men økningen i antall siteringer er likevel så stor at man ikke kan se bort fra den. Her vil jeg trekke frem emne kategorien MED som har en vesentlig økning hos alle institusjonene, og UiO spesielt med en økning på hele 1028,3 %. Økningen i antall siteringer for UiO er også høy for de andre emne kategoriene, med henholdsvis 245,0 % i MAT og 247,6 % i SAMF. Igjen er økningen av antall siteringer i hovedsak vesentlig, og kan ikke alene forklares med fordelingen av artikler og lengde på siteringsperiodene.

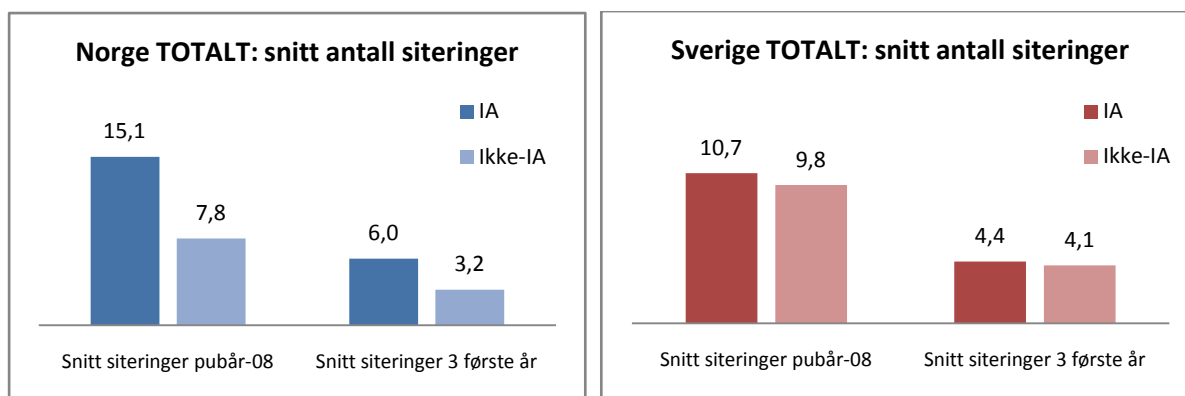
### 5.1.2 Endringer i antall siteringer for artikler parallellpublisert i IA eller ikke

I denne analysen tok jeg utgangspunkt i det samme artikkelutvalget fra ISI, men plukket ut artiklene med match mellom ISI og NORA/DiVA, det vil si at artiklene finnes begge steder. Videre foretok jeg et tilfeldig plukk av samme antall artikler som ikke er parallellpublisert tilknyttet de samme forfatterne, dvs. at artiklene finnes kun i ISI. Jeg prøvde å plukke artikler

publisert innen samme måned og årstall for å få et mest mulig likt grunnlag for sammenligning, men dersom det ikke var mulig ble den artikkelen som var publisert nærmest i tid brukt. De to grupperingene definert som ”IA artikler”, dvs. artikler som både finnes i ISI og NORA/DiVA, og ”ikke-IA artikler”, dvs. de artiklene jeg har plukket fra ISI, som kun finnes i ISI og som ikke er parallellpublisert, blir sammenlignet og målt mot hverandre.

Jeg summerte antall siteringer totalt for IA artiklene og ikke-IA artiklene for periodene før og etter innføring av IA, beregnet totaler og snitt i antall og prosent pr forsker. I tillegg ble dette summert på nivåene emnekategori, institusjon og land. De akkumulerte tallene gjorde det mulig å sammenligne og se på eventuell utvikling for grupper av forskere innen emnekategoriene og institusjonene i Norge opp mot hva som skjer i Sverige.

Søylene i Figur 14 viser snitt antall siteringer for IA artikler sammenlignet med ikke-IA artikler. De to første søylene viser snitt siteringer fra artiklenes publiseringsår frem til og med 2008. De to siste søylene viser snitt siteringer i artiklenes 3 første år etter publisering. Snittet er beregnet i forhold til antall forfattere som har artikler med match og ikke i forhold til antall artikler. For Norge er antall forfattere i analyseutvalget 68, mens det for Sverige er 160 forfattere.

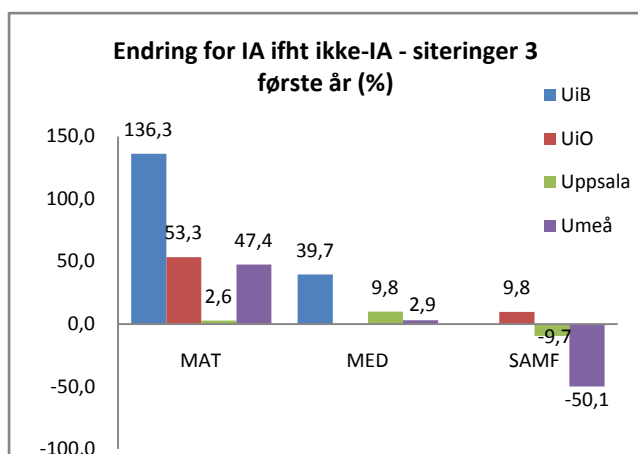


**Figur 14 - snitt antall siteringer pr forfatter for IA artikler sammenlignet med ikke-IA artikler for tidsperioden publiseringsår-08 og de 3 første årene etter publisering.**

Både ved å se på snitt antall siteringer for hittil levetid for artiklene, dvs. fra publiseringsår til og med 2008, og isolert på de tre første årene etter publisering viser søylene en økning i antall siteringer for IA artikler sammenlignet med ikke-IA artikler. Dette gjelder både for Norge og Sverige, og en ser at økningen er noe større i Norge enn i Sverige.

Ser man på de første to søylene som viser snitt siteringer fra artiklenes publiseringsår frem til 2008 har norske artikler et snitt på 15,1 for IA artikler mot 7,8 for ikke-IA artikler. Søylene for snitt antall siteringer de 3 første årene etter publisering viser også en økning hvor norske IA artikler har et snitt på 6,0 mot 3,2 for ikke-IA artikler. Økningen i snitt antall siteringer er på henholdsvis 94,3 % og 87,0 % for de to beregningene. De samme tallene fra Sverige er ikke like høye. Her er snittet på antall siteringer for perioden publiseringsår-08 for IA artikler er på 10,7 mot 9,8 for ikke-IA artikler. Når det gjelder snittet for siteringer i de 3 første leveårene er dette på 4,4 for IA artikler mot 4,1 for ikke-IA artikler. Forskjellen her er på henholdsvis 8,6 % og 7,1 %, men fortsatt en økning for IA artikler i forhold til ikke-IA artikler.

For å finne en mulig forklaring på den store forskjellen mellom Norge og Sverige må man se på detaljene knyttet til emnekategoriene og de enkelte institusjonene. Beregningene for snitt antall siteringer for artiklenes 3 første leveår gir et ensartet og sammenlignbart datagrunnlag siden jeg bruker de samme 3 årene for begge kategorier, dvs. de 3 første årene både for IA og

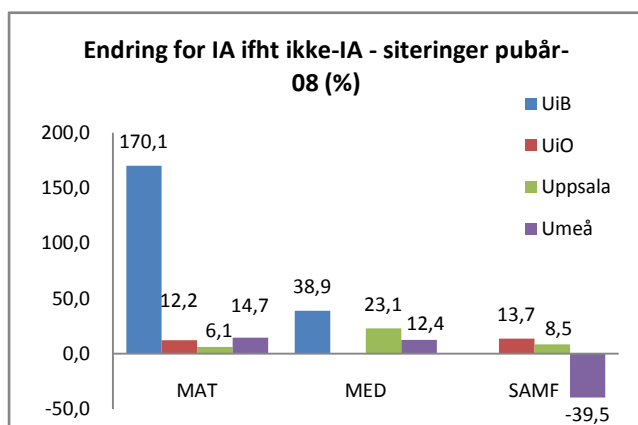


Figur 15 - endring i snitt antall siteringer pr forfatter for IA artikler sammenlignet med ikke-IA artikler for tidsperioden de 3 første årene etter publisering.

ikke-IA artikler. Søylene i Figur 15 viser at det er en økning, om litt variabel, for alle emnekategoriene og alle institusjonene med unntak av emnekategori SAMF. Her er det flere siteringer for ikke-IA artikler enn for IA artikler både for Uppsala og Umeå. Når det gjelder Umeå har IA artiklene halvparten så mange siteringer som ikke-IA artiklene (-50,1 %). For Uppsala er forskjellen noe mindre med en reduksjon på 9,7 %.

Videre kan man se at for emnekategori MED er forskjellen også liten mellom IA og ikke-IA artikler for Uppsala og Umeå, mens den i tillegg er liten for emnekategori MAT for Uppsala. Her har Umeå en økning i snitt antall siteringer for IA artikler på 47,4 % sammenlignet med snitt antall siteringer for ikke-IA artikler. Jeg gjør oppmerksom på at UiB ikke har artikler med match i emnekategori SAMF, og at UiO ikke har artikler med match i emnekategori MED. Dette er tatt hensyn til i beregningene, og er derfor heller ikke med i figurene over eller i totaltallene for Norge.

Det som er interessant er at for emnekategori MAT har UiB en økning på hele 136,3 % for IA artikler i forhold til ikke-IA artikler. UiO og Umeå har også en sterk økning i emnekategori MAT på henholdsvis 53,3 og 47,4 %. Ellers har som nevnt Umeå kun en liten økning i emnekategori MED på 2,9 %, men i tillegg en reduksjon på hele 50,1 % i kategori SAMF. Hva som er årsaken til disse svingningene er vanskelig å si. Det kan være relatert til at det er en viss forskjell på antall artikler med match i de forskjellige emnekategoriene. For eksempel har emnekategori SAMF for Umeå bare halvparten av antall matchende artikler i forhold til MAT, og ca 2/3 av antall matchende artikler i forhold til MED. Dette gjelder også for Uppsala hvor emnekategori SAMF har halvparten av antall matchende artikler i forhold til MAT, men der er det en positiv utvikling i antall siteringer med 9,8 %.



Figur 16 - endringer i snitt antall siteringer pr forfatter for IA artikler sammenlignet med ikke-IA artikler summert for tidsperioden fra publiseringsår til og med 2008.

Figur 16 viser endringer for IA artikler sammenlignet med ikke-IA artikler basert på snitt antall siteringer for artiklenes levetid, dvs. fra artiklenes publiseringsår til og med 2008. Ser man denne analysen i forhold til den forrige, som tok utgangspunkt i de 3 første årene etter publisering av artiklene, viser også denne en økning i antall siteringer for IA artikler i forhold til ikke-IA artikler, igjen med unntak av emnekategori SAMF for Umeå.

Beregningen på snitt antall siteringer for tidsperioden publiseringsår til og med 2008 er interessant, men fordi noen artikler i utvalget ble publisert tidlig i perioden, mens andre ble publisert sent, blir tallmaterialet noe mer sprikende og skjevt fordelt. Likevel bekrefter denne beregningen økningen i snitt antall siteringer for IA artikler sammenlignet med ikke-IA artikler.

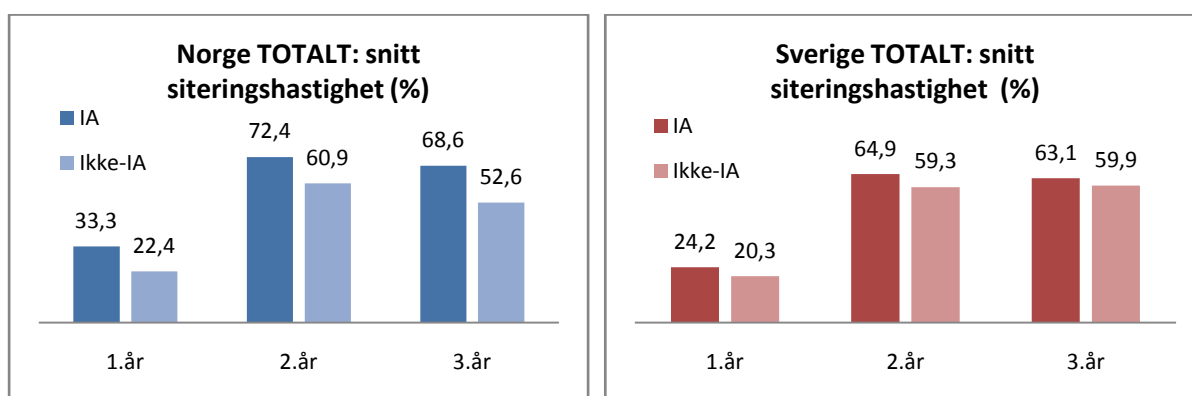
## 5.2 Siteringshastighet

For å svare på det andre forskningsspørsmålet; om artikler som er parallellpublisert i IA blir raskere sitert sett i forhold til artikler som kun er publisert i tradisjonelle kanaler, foretok jeg tre analyser hvor jeg så på siteringshastigheten fra ulike vinkler. Alle analysene sammenligner artikler som er parallellpublisert med artikler som ikke er det. Først så jeg på hvor raskt siteringene startet og fordelingen av siteringene de 3 første årene fra og med publiseringsåret. Deretter så jeg på om innføring av IA medførte en ”ny giv”, dvs. om artikler publisert før innføring av IA fikk en økning i antall siteringer etter innføringen, dvs. om antall siteringer ble påvirket og fikk et oppsving. Den siste analysen ga en tilsvarende oversikt som ”ny-giv” analysen, men her er siteringene årsfordelt fra publiseringsåret frem til og med 2008. Resultatene er presentert og diskutert for hver av analysene i de neste underkapitlene.

### 5.2.1 Endringer i siteringshastighet for artikler parallellpublisert i IA eller ikke

For å finne siteringshastigheten tok jeg utgangspunkt i de samme to grupperingene med artikler som i kapittel 5.1.2, dvs. utvalget med IA og ikke-IA artikler. Jeg så på hvor mange IA og ikke-IA artikler som ble sitert de tre første leveårene fra og med publiseringsåret, beregnet snitt pr år og sammenlignet de to grupperingene.

Søylene i Figur 17 viser siteringshastigheten, dvs. hvor stor prosent av IA/ikke-IA artiklene som blir sitert hvert år de 3 første årene etter publisering.



Figur 17 – siteringshastigheten viser hvor stor prosent av IA/ikke-IA artiklene som blir sitert hvert år de 3 første årene etter publisering.

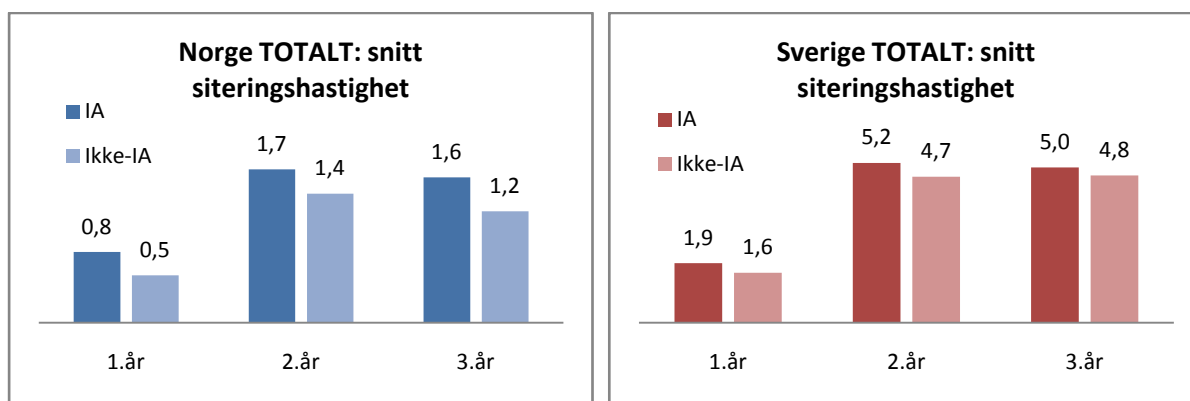
Tallene både fra Norge og Sverige viser at en høyere andel av IA artiklene blir sitert hvert år de første 3 årene etter publisering enn ikke-IA artiklene. For Norges vedkommende betyr det at totalt 33,3 % av de norske IA artiklene blir sitert allerede det første året, mens kun 22,4 %

av ikke-IA artiklene siteres. Andre året øker andelen, og 72,4 % av IA artiklene siteres mot 60,9 % av ikke-IA artiklene. Tredje året synker dette noe, men fortsatt har IA artikler en høyere siteringsandel hvor 68,6 % siteres mot 52,6 % av ikke-IA artiklene. Dette viser at sjansen for å bli sitert er høyere hvert av de 3 første årene etter publisering dersom artikkelen er parallellpublisert enn om den ikke er det. Økningen i siteringshastighet for IA artikler i forhold til ikke-IA artikler er på hele 48,6 % første år, 18,9 % andre år og 30,5 % tredje år.

Tallene fra Sverige viser samme tendens, men ligger noe lavere enn Norge, og det er jevnere mellom IA artikler og ikke-IA artikler. Totalt blir 24,2 % av de svenske artiklene sitert allerede det 1. året, mens 20,3 % av ikke-IA artiklene siteres. Andre året øker siteringsandelen og 64,9 % av IA artiklene siteres mot 59,3 % av ikke-IA artiklene. Tredje året synker andelen for IA artikler noe, mens den faktisk øker for ikke-IA artikler, men fortsatt har IA artikler en høyere siteringsandel med 63,1 % mot 59,9 % av ikke-IA artiklene. Dette viser også at i Sverige er sjansen for å bli sitert høyere hvert av de 3 første årene etter publisering dersom artikkelen er parallellpublisert enn om den ikke er det. Økningen i siteringshastighet for IA artikler i forhold til ikke-IA artikler er lavere enn i Norge, men er fortsatt på 18,8 % første år, 9,5 % andre år og 5,4 % tredje år.

En detalj, men interessant, er at om man ser bak prosentene i Figur 17, og henter frem antall siteringer for IA og ikke-IA artikler i Norge og Sverige, viser snitt antall siteringer de første 3 årene etter publisering at svenske artikler i snitt har et høyere faktisk antall siteringer enn norske artikler.

Søylene i Figur 18 viser snitt antall siteringer for IA og ikke-IA artikler hvert år de 3 første årene etter publisering. Snittet er beregnet i forhold til antall forfattere som har artikler med match, og ikke i forhold til antall artikler som siteres hvert år som den prosentvise siteringshastigheten i Figur 17 baserer seg på.



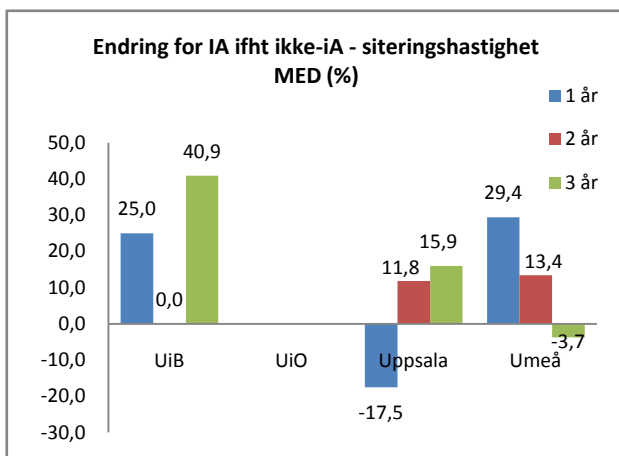
**Figur 18 – snitt siteringshastighet viser snitt antall siteringer pr forfatter for IA/ikke-IA artikler hvert år de 3 første årene etter publisering.**

At Sverige ligger over Norge i snitt antall siteringer pr forfatter hvert år de 3 første årene etter publisering kan delvis forklares med et høyere antall artikler i det svenske artikkelutvalget enn hva det norske utvalget har. Dette gjelder for både IA og ikke-IA artikler. I tillegg har det svenske utvalget flere artikler med match mellom ISI og DiVA enn hva jeg finner i det norske utvalget. Jeg viser til detaljer om forskjellene i kapittel 4.1. Likevel opererer jeg med snittberegninger for nettopp å utjevne disse forskjellene og for å få tallene sammenlignbare. Tallene konkluderer med at svenske artikler i snitt siteres oftere enn norske. Årsakene kan være mange, men det har klart en sammenheng med at ISI har indeksert flere artikler hvor det er svenske forfattere/medforfattere enn artikler med norske forfattere. Det kan igjen være et resultat av at svenske forfattere publiserer mer enn norske, eller at de publiserer mer i internasjonale enn nasjonale/nordiske tidsskrifter enn norske forfattere.

Ser man nærmere på tallene for institusjoner og emne kategorier i forhold til siteringshastighet, viser også tallene på institusjons- og emne kategorinivå at det er en høyere andel av IA artiklene som blir sitert hvert år de første 3 årene etter publisering enn ikke-IA artiklene. Her er det likevel noen interessante variasjoner – viser til Figur 19, Figur 20 og Figur 21 nedenfor. Jeg gjør oppmerksom på at UIB ikke har artikler med match med ISI i emne kategorien SAMF, og UIO ikke har artikler med match med ISI i emne kategorien MED. Dette er tatt hensyn til i beregningene, og derfor heller ikke med i totaltallene eller grafene for Norge.

Emne kategorien MED skiller seg ut – se Figur 19. For Uppsala varierer forskjellen i siteringshastigheten for IA artikler i forhold til ikke-IA artikler fra å være negativ, (-17,5 %) første år, som betyr at ikke-IA artikler var mer sitert enn IA artikler, til å bli positiv og øke

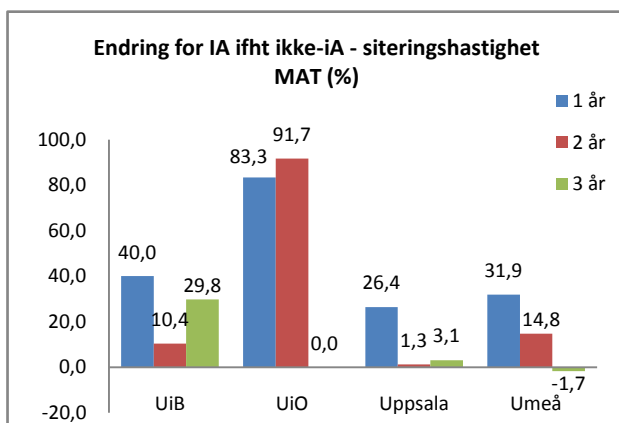




Figur 19 – endringer i siteringshastighet hvert år de 3 første årene etter publisering for IA artikler i forhold til ikke-IA artikler innen emnekategorien medisin (MED).

uttrekk kun 2 medisinske fagfelleverderte tidsskriftsartikler registrert i NORA, og disse hadde ingen match med artikler i ISI, og dermed er det ingen data for UiO her i Figur 19. For UiB er det en positiv økning i siterings-hastigheten i to av de tre første årene etter publisering i IA artiklers favør i forhold til ikke-IA artikler med henholdsvis 25,0 % første år, ingen forskjell andre år (0 %) og 40,9 % tredje år.

Hva som er årsaken til disse variasjonene er vanskelig å si på bakgrunn av det tallmateriale jeg har. Antall artikler med match mellom ISI og DiVA, samt antall forfattere tilknyttet disse artiklene er nesten identisk hos Uppsala og Umeå, så forskjellen ligger i antall siteringer. Siden jeg har plukket ut forfatternavn basert på rangerte lister i forhold til hvor mange artikler de er registrert med i NORA/DiVA kan jeg heller ikke spekulere i om populariteten i temaene i artiklene, eller om noen av forfatterne er mer kjente, og dermed mer sitert, enn andre har innvirkning på antall siteringer.

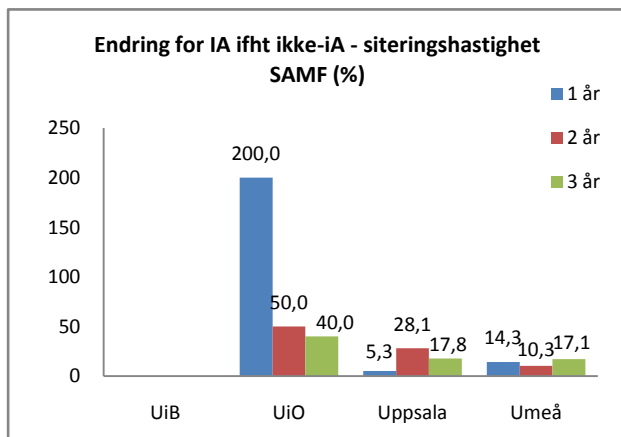


Figur 20 - endringer i siteringshastighet hvert år de 3 første årene etter publisering for IA artikler i forhold til ikke-IA artikler innen emnekategorien matematikk og naturvitenskap (MAT).

annet og tredje år med henholdsvis 11,8 og 15,9 %. For Umeå er det helt motsatt. Her starter forskjellen i siteringshastighet mellom IA/ikke-IA artikler i IA artiklers favør og snues deretter på hodet, slik at i det tredje året etter publisering er andelen høyere for ikke-IA artikler. IA artikler har altså en økning i forhold til ikke-IA artikler på 29,4 % første år, 13,4 % andre år og en reduksjon på 3,7 % tredje år. Når det gjelder UiO var det på tidspunktet for

For emnekategorien MAT ser det mer ”normalt” ut i forhold til totaltallene – vist her i Figur 20. Med ”normalt” mener jeg at tallene bekrefter analyseresultatene hittil, dvs. en trend som viser en økning i siteringshastighet for IA artikler i forhold til ikke-IA artikler. Det som er interessant er at for nesten alle institusjonene er økningen i hvor mange artikler som

siteres sett IA artikler i forhold til ikke-IA artikler størst det første året og deretter noe mindre i andre og tredje år. Den samme tendensen viser tallene i emnekategori SAMF med noen variasjoner – se Figur 21.



Figur 21 - endringer i siteringshastighet hvert år de 3 første årene etter publisering for IA artikler i forhold til ikke-IA artikler innen emnekategori samfunnsfag (SAMF).

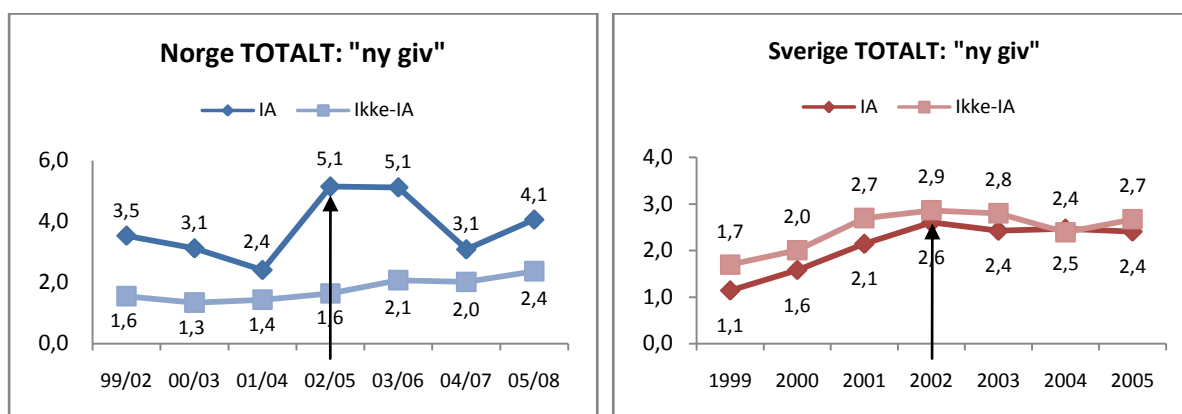
Dette betyr at mitt tallmateriale gir en indikasjon på at artikler som er parallellpublisert i IA siteres raskere enn artikler som ikke er parallellpublisert. Når IA artikler siteres oftere allerede første år etter publisering indikerer det at IA artikler ”oppdages” raskere enn ikke-IA artikler. Igjen kan jeg ikke tillegge hele forklaringen for den raskere siteringshastigheten det faktum at artiklene er parallellpublisert i IA, det kan være andre faktorer som påvirker, men tendensen er klar.

### 5.2.2 ”Ny giv” – oppsving i antall siteringer for ”gamle” artikler

En annen vinkling i forhold til siteringshastighet og økning i antall siteringer var å se på om innføring av IA medførte en ”ny giv”, eller et oppsving i antall siteringer. Her var utgangspunktet det samme artikkelutvalget som i den foregående analysen, dvs. artikler med match mellom ISI og NORA/DiVA, og tilsvarende antall ikke-IA artikler. For å kunne sjekke endringer og eventuell økning i antall siteringer analyserte jeg kun de artiklene som var publisert minimum 3 år før innføring av IA, og brukte et tilsvarende antall ikke-IA artikler med samme publiseringsår som kontrollgruppe. Jeg undersøkte en begrenset periode før og etter innføringstidspunktet for IA, for å se om innføringen ga forskjeller for IA/ikke-IA artikler i forhold til endringer i antall siteringer og eventuelt et oppsving/hopp i siteringene for disse ”gamle” artiklene. Måleperioden jeg satt var 3 år før innføring av IA, selve innføringsåret og 3 år etter innføring av IA. For UiB som hadde startåret sitt i 2005 hentet jeg alle IA/ikke-IA artiklene publisert i 2002 og tidligere, slik at jeg kunne undersøke siteringene i intervallet 02-08 (før: 02-04, IAår:05 og etter:06-08). For UiO, Uppsala og Umeå som alle hadde oppstart av IA i 2002 hentet jeg artikler publisert i 1999 og tidligere for å kunne undersøke et tilsvarende intervall for 99-05 (før: 99-01, IAår:02 og etter:03-05).

Denne undersøkelsen reduserte utvalget mitt betraktelig. For Norges vedkommende ble utvalget redusert fra totalt 2111 artikler funnet i ISI tilknyttet de 118 forfatternavnene jeg hadde plukket etter artikkelsøket i NORA. Dette ble ytterligere redusert til 156 artikler med match mellom ISI og NORA, og hvor kun 42 artikler var publisert før henholdsvis 2002 for UiB og 1999 for UiO. Disse artiklene var tilknyttet 17 av det opprinnelige utvalget på 118 forfatternavn. For Sverige ble utvalget redusert fra totalt 8.999 artikler funnet i ISI tilknyttet de 173 forfatternavnene jeg hadde plukket etter artikkelsøket i DiVA. Dette ble så redusert til 1278 artikler med match, og deretter til kun 315 artikler som var publisert før 1999 tilknyttet 48 forfatternavn. Dette er ikke et stort analysegrunnlag. Landene sett under ett utgjør kun 24,9 % av antall artikler i utvalget med match mellom ISI og NORA/DiVA (357 mot 1434), og faktisk kun 3,2 % av brutto-/nettogrunnlaget som er det totale antall artikler hentet fra ISI for valgte forfattere og emne kategorier (357 mot 11.110).

Figur 22 viser "Ny giv", dvs. en oversikt over snitt antall siteringer for en 7 års periode for IA og ikke-IA artikler. Perioden på 7 år strekker seg som nevnt tidligere over intervallet 3 år før innføring av IA, selve innføringsåret og 3 år etter innføring av IA. Alle artiklene i analyseutvalget er publisert første år i perioden eller tidligere. Teksten på x-aksen for Norge er ikke rene årstall, fordi grafen viser akkumulerte verdier fra UiB og UiO som hadde oppstart av sine IA (BORA og DUO) i henholdsvis 2005 og 2002. Tidsintervallet i analysen er dermed 2002-2008 for UiB og 1999-2005 for UiO. Teksten gjenspeiler begge intervallene.



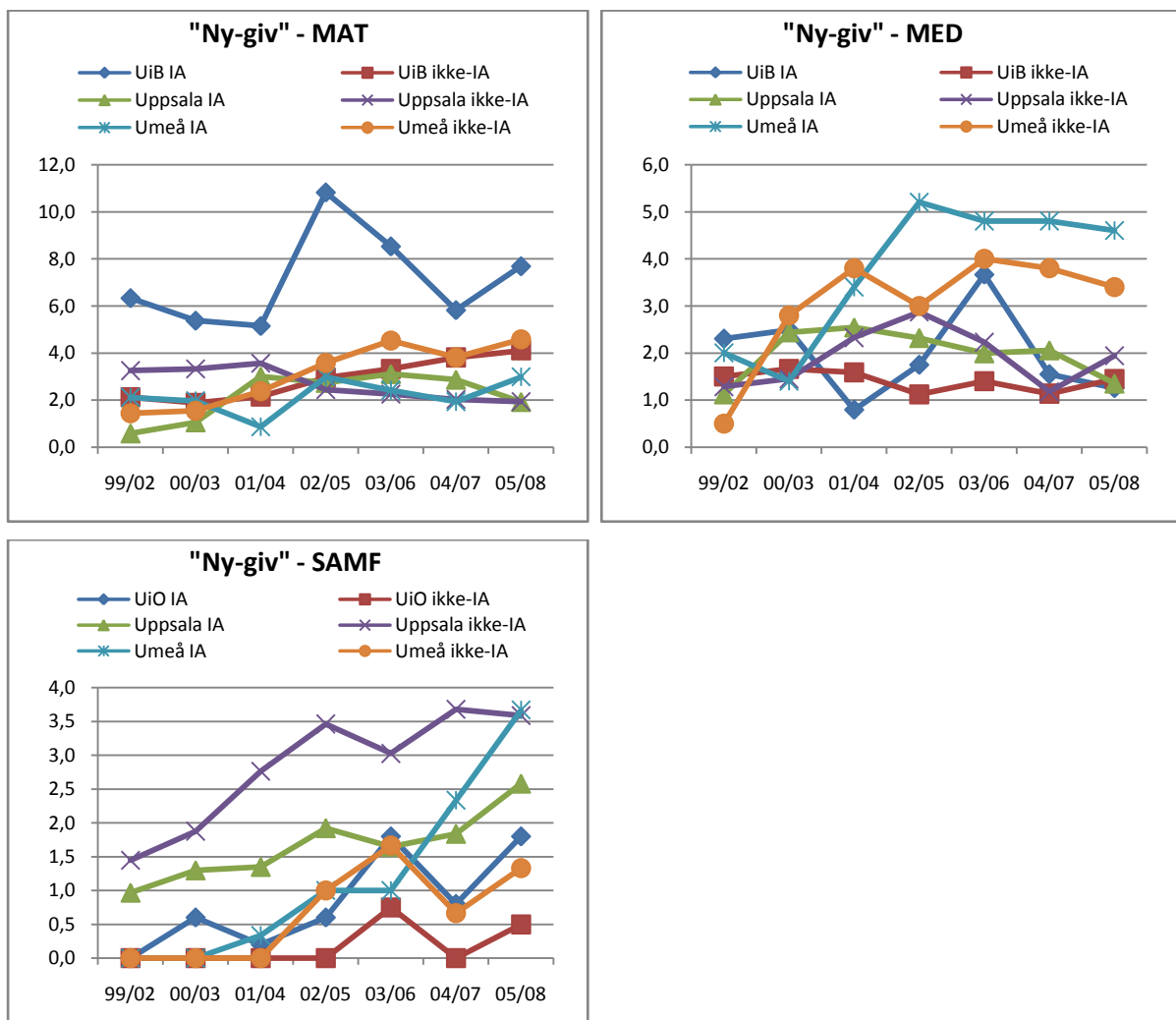
Figur 22 - "ny giv" viser snitt antall siteringer over en 7 års periode for IA/ikke-IA artikler.

Det første som slår en når man ser grafene i Figur 22 er forskjellen mellom Norge og Sverige, og det andre som slår en er variasjonen i siteringene for IA artikler for Norge i måleperioden. For å ta det første først; Kurvene for Sverige viser at ikke-IA artikler er jevnt over mer sitert enn IA artikler, men begge grupperingene har en økning i snitt antall siteringer i

måleperioden, og IA artikler mer enn ikke-IA artikler. Økningen er på 111,2 % for IA artikler (fra 1,1 til 2,4) og på 57,4 % for ikke-IA artikler (fra 1,7 til 2,7). Det er få siteringer i starten og deretter en liten økning år for år. Siteringene nådde en "topp" i 2002 med et snitt på henholdsvis 2,9 og 2,6 siteringer for ikke-IA og IA artikler før en viss utflating. IA artiklene ligger lavere i snitt antall siteringer i hele perioden, men den lille nedgangen som kommer til uttrykk i 2003 med et snitt på 2,4 fortsetter ikke, men derimot dreier og øker igjen til et snitt på 2,5 i 2004 før snittet går tilbake til 2003-nivå i 2005. For ikke-IA artikler skjer noenlunde det samme. Om de små retningsendringene skyldes innføring av IA ved institusjonene er vanskelig å konstatere og umulig å bekrefte, men det at snittet flater raskere ut for IA artikler sammenlignet med ikke-IA etter "topp"-året i 2002 kan kanskje sees i sammenheng med resultatet av siteringshastighetsanalysen. Der fant jeg at sjansen for å bli sitert var høyere hvert av de 3 første årene etter publisering dersom artikkelen var parallellpublisert enn om den ikke var det – ref. kapittel 5.2.1 Endringer i siteringshastighet for artikler parallellpublisert i IA eller ikke.

Går man over til Norge og ser på fordelingen av snitt antall siteringer pr år over "ny-giv"-perioden så følger utviklingen i Norge noe av den samme fordelingen som Sverige til tross for større svingninger. Forskjellen ligger at det for Norge er IA artiklene som har det høyeste snittet sammenlignet med ikke-IA artikler, at det er større avstand mellom snitt antall siteringer for IA og ikke-IA artikler, og at det for IA artikler er en større variasjon i snittet år for år, dvs. ikke like jevn kurve som Sverige. At snitt antall siteringer for IA artikler ligger høyere enn ikke-IA artikler også i perioden før innføring av IA kan være tilfeldig, og artikkelutvalget gir ingen svar på om det er noe spesielt med disse artiklene sammenlignet med ikke-IA artiklene. Jeg nevnte at det ville være av interesse for min undersøkelse å vite om dette var høyt anerkjente artikler plukket ut spesielt for etterregistrering, eller om det var artikler som forskerne selv har valgt å etterregistrere av egeninteresse, men fordi jeg ikke har gått inn i problemstillingen kan jeg bare spekulere. Det interessante også for Norge er at både IA og ikke-IA artikler har en retningsendring et par år etter tidspunktet for innføring av IA på samme måte som for Sverige. Snittet på antall siteringer øker igjen etter en nedgang, og økningen er høyere for IA artikler sammenlignet med ikke-IA artikler i samme periode. Snitt antall siteringer øker med henholdsvis 31,5 og 17,4 % for IA og ikke-IA artikler mellom 04/07 og 05/08. Om det er innføringen av IA som har vært påvirkningsfaktor for å snu nedgangstrenden som startet etter 03/06 er like vanskelig å konstatere for Norge som det er for Sverige.

Når man ser på "ny-giv"-fenomenet for hver av emnekategoriene i Figur 23 er det for Sverige kun kategorien SAMF som bekrefter oppsvinget i snitt antall siteringer et par år etter innføring av IA. Både emnekategoriene MAT og MED har i hovedsak en videre nedgang etter "toppen", og ikke en positiv retningsendring i årene etter 2002. Unntaket er at det for Umeå også er en retningsendring i emnekategori MAT som viser en ny økning i antall siteringer. Dette gjelder riktignok både IA/ikke-IA artikler. Den er ikke godt synlig i grafen for emnekategori MAT i Figur 23, men økningen i snitt antall siteringer etter siste retningsendring i 04/07 er høyere enn nedgangen fra "toppårene" for hver av grupperingene IA/ikke-IA artikler. Beregninger viser at økningen er høyere for IA artikler i forhold til ikke-IA artikler, selv om ikke-IA artikler relativt sett fortsatt ligger høyere i snitt antall siteringer enn IA artikler.

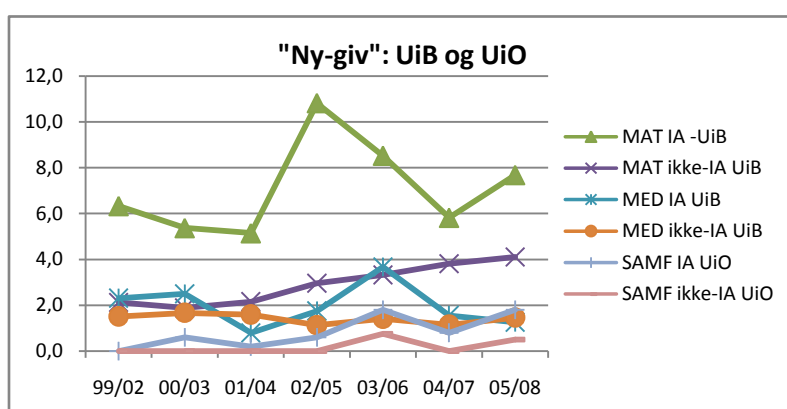


Figur 23 - "ny giv" viser snitt antall siteringer over en 7 års periode for IA artikler pr emnekategori

Begrunnelsen for at UiB og UiO ikke er vist i alle grafene i Figur 23 er som tidligere nevnt at UiB ikke har artikler med match med ISI i emnekategori SAMF, og UiO ikke har artikler

med match med ISI i emnekategoriene MED. I tillegg var de artiklene UiO hadde match i kategorien MAT for ”ferske” i forhold til denne analysen, dvs. artiklene er publisert etter 1999. Dette er tatt hensyn til i beregningene, og derfor heller ikke med i totaltallene for Norge eller i grafene over.

Ser man nærmere på detaljene for emnekategoriene og institusjonene isolert for Norge gjengitt i Figur 24 viser alle ”ny-giv”-kurvene en positiv retningsendring et par år etter innføringstidspunktet for IA med unntak av emnekategoriene MED, som har den karakteristiske nedgangen og utflatingen.



Figur 24 – utviklingen av snitt antall siteringer i ”ny giv”-perioden for IA/ikke-IA artikler hos universitetene i Bergen og Oslo (UiB og UiO)

Siden totaltallene for Norge i Figur 24 baserer seg på akkumulerte tall fra ulike tidsintervaller for UiO og UiB (henholdsvis 99-05 for UiO og 02-08 for UiB), fremkommer det ikke like tydelig i denne figuren at snitt antall siteringer øker et par år

etter innføring av IA for alle emnekategoriene. Økningen gjelder både IA og ikke-IA artikler og for alle emnekategoriene unntatt kategorien MED hvor UiB har en nedgang i antall siteringer for IA artikler, men en økning for ikke-IA artikler. Dette er tydeligere visualisert i Figur 23.

En annen viktig observasjon for Norge er ulikhetene i snittverdiene. Snitt antall siteringer for emne-kategoriene MAT og MED, dvs. verdier for UiB, ligger høyere enn UiOs snittverdier for emnekategoriene SAMF. Noe av forklaringen kan være at det dreier seg om ulike emne-kategorier, og at det dermed er naturlig med ulike snittverdier. En annen forklaring kan ligge i selve datagrunnlaget. Selv om observasjonene er små har jeg et datagrunnlag for UiB hvor antall artikler med match publisert tidligst 3 år før innføring av IA er 7,4 ganger større enn tilsvarende datagrunnlag for UiO (37 artikler mot 5). Det er fire ganger så mange artikler i grunnlaget mitt for UiBs MED som i UiOs SAMF (22 artikler mot 5), og det er 3 ganger så mange artikler i grunnlaget for UiBs MAT i forhold til UiOs SAMF (15 artikler mot 5). Til

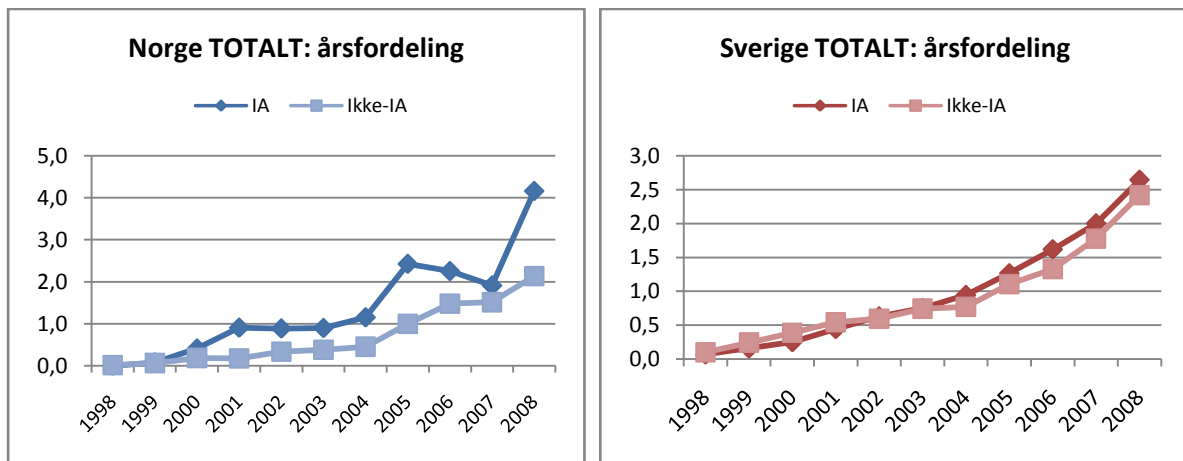
tross for at jeg opererer med snittberegninger er det klart at med et større og jevnere datagrunnlag kunne tallene sett annerledes ut. Det som styrker de resultatene jeg likevel har funnet er at antall artikler i datagrunnlaget for UiBs emnekategori MED er høyere enn for kategorien MAT, og likevel har emnekategori MAT jevnt over et mye høyere snitt enn kategorien MED.

Kurvene i Figur 22, Figur 23 og Figur 24 tegner et bilde av at det skjer et hopp/en økning i antall snitt siteringer et par år etter tidspunktet for innføring av IA. Som nevnt er det ikke mye data jeg baserer disse fordelingen på, men jeg mener det er nok til å indikere en utviklingstendens. Tallmaterialet mitt viser en svak positiv retningsendring for nesten alle emne kategorier og institusjoner, spesielt i Norge. De endringen og økningene jeg ser tendensen av i grafene kan skyldes innføring av IA, men også andre faktorer kan ha spilt en rolle. Både IA og Open Access som fenomen er nye, og det vil nok for norske forhold ta enda noe tid før dette er allment kjent og gi en mer synlig positiv effekt i antall siteringer.

### **5.2.3 Årsfordeling av siteringer i perioden 1998-2008**

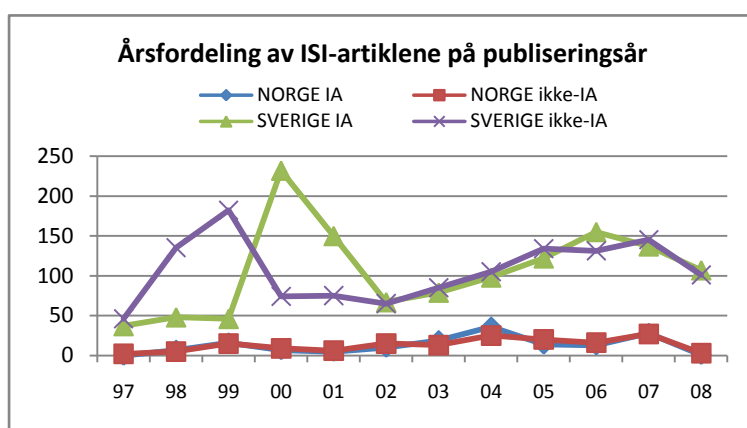
Siden ”ny giv”-analysen min kun baserer seg på artikler publisert senest 3 år før innføring av IA, fikk jeg et veldig lite datamateriale å analysere. For å bruke en annen vinkling på siterings hastighet og snitt antall siteringer over en tidsperiode har jeg derfor i tillegg sett på utviklingen av snitt antall siteringer pr år i tidsintervallet 1998-2008 for både IA og ikke-IA artikler. Hele måleperioden for undersøkelsen er 1997-2008, men fordi siteringer ofte ikke starter før et par år etter publisering har jeg valgt måleperioden 1998-2008 for denne analysen. Minner også om at IA artikler er artikler som både finnes i ISI og NORA/DiVA, og ikke-IA artikler er de artiklene som kun finnes i ISI og som ikke er parallellpublisert. Ved å bruke samme datagrunnlag som de øvrige analysene av endringer i antall siteringer for artikler parallellpublisert i IA eller ikke fikk jeg et årsfordelt siteringssnitt over en lengre periode enn ”ny giv”-analysen. Utgangspunktet er snitt antall siteringer for både IA og ikke-IA artikler fordelt pr år i tidsintervallet. Snittet baserer seg på et aggregat av snitt antall siteringer pr. artikkel pr. forfatter innen hvert emne.

Kurvene i Figur 25 viser snitt antall siteringer for IA/ikke-IA artikler pr år i tidsintervallet 1998-2008.



Figur 25 - årsfordeling av snitt antall siteringer for IA/ikke-IA artikler i tidsintervallet 1998-2008

Siteringskurvene både for Norge og Sverige er som forventet med en oppadstigende kurve. I og med at datagrunnlaget baserer seg på alle artiklene i utvalget for de to kategoriene IA/ikke-IA artikler, uavhengig av publiseringsår, betyr det at det i utvalget finnes artikler publisert allerede i 1997 med siteringsdata for hele tidsperioden, og det finnes artikler av nyere dato som er publisert helt i slutten av tidsperioden med siteringsdata for få år. Alle siteringer er summert og snittfordelt pr år. Mange artikler publisert tidlig i perioden betyr flere år med siteringsdata, og mange artikler publisert sent i perioden betyr færre år med siteringsdata. Kurvene er likevel oppadstigende fordi alle artiklene i utvalget er publisert i tidsintervallet og de fleste er på vei til å bli ”oppdaget”, dvs. antall siteringer pr artikkel er på vei opp.



Figur 26 - artiklene fra ISI fordelt pr publiseringsår

For Norges vedkommende er fordelingen av artikler for IA og ikke-IA artikler jevn over alle årene i tidsintervallet. Kurvene i Figur 26 viser at publiseringsårene stort sett er overlappende i hele intervallet ved at linjene nesten dekker hverandre. Antall artikler fra Sverige er en god del høyere enn antallet fra Norge

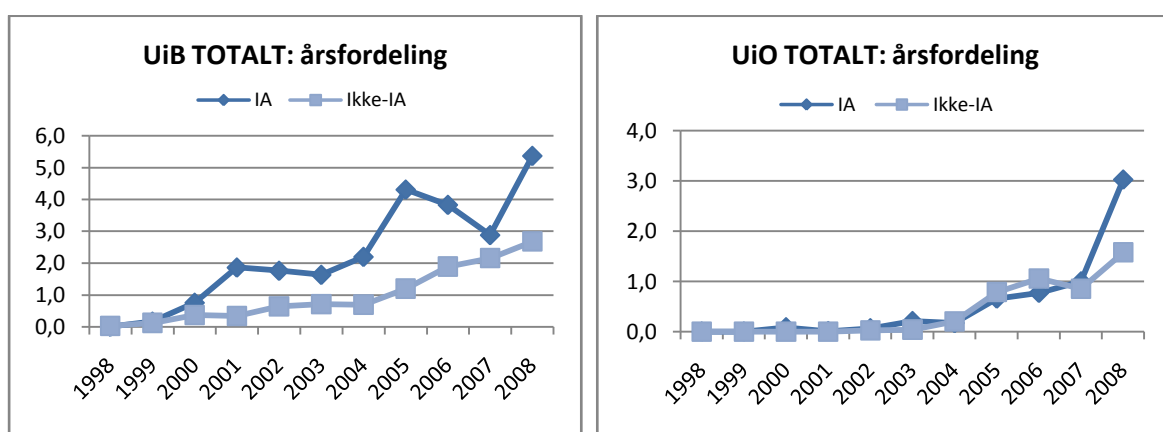
(Sverige har 1278 og Norge har 156 artikler med match). Dette er lett å se ved at kurvene for Sverige ligger over de norske. Kurvene for Sverige viser at det er en parallellforskyvning med et høyere antall ikke-IA artikler fra perioden 1998-1999 og et tilsvarende høyt antall IA artikler fra 2000-2001. Et høyt antall artikler i starten av tidsintervallet betyr siteringsdata



over mange år, mens Norge som har sin ”topp” antall artikler (liten, men dog en topp) i 2004 har relativt flere artikler med ”oppadgående” antall siteringer i slutten av tidsintervallet. Ujevnt tilsig av artikler i ISI kan bety ujevnt tilsig av antall siteringer, slik at i en analyse hvor siteringer diskuteres vil slike topper kunne forårsake støy i analyseresultatene. Jeg har i mine analyser ikke justert for disse ujevnheterne. Artiklenes årsfordeling blir diskutert mer i detalj senere i dette kapittelet når analysene for det enkelte land presenteres.

Ser man tilbake på kurvene i Figur 25 som viser årsfordelinger pr land, ser man at kurvene som viser IA/ikke-IA artikler ligger forholdsvis parallelt med hverandre, mer for Sverige enn for Norge. Det som er interessant er at kurven for IA artikler ligger noe over kurven for ikke-IA artikler som bekrefter de tidligere analysene, men forskjellen er så liten at jeg skal være forsiktig med å vektlegge det for mye. Det som i tillegg er interessant når man ser på kurvene for Norge er retningsendringene som skjer i 2004 og på ny i 2007. DUO ble innført i 2002 ved UiO, mens BORA ble innført i 2005 ved UiB. Retningsendringene skjer altså et par år etter innføring av IA, og man vil kunne anta at innføring av IA har hatt en positiv effekt, selv om det er vanskelig å konstatere.

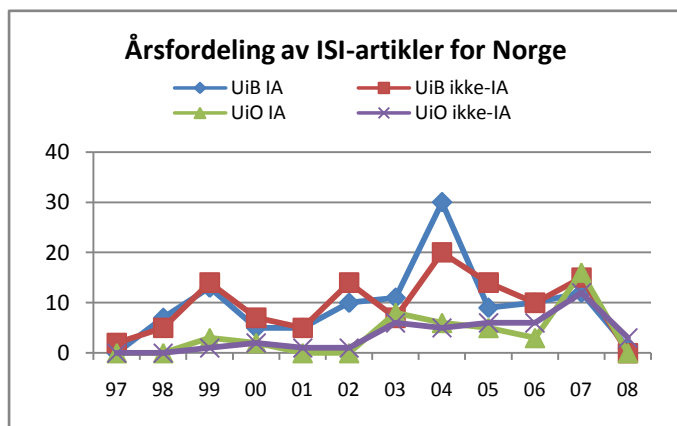
Forklaringen er kanskje for enkel, for når man trekker frem grafer som viser snitt antall siteringer pr år for hver av institusjonene, er det fortsatt de samme retningsendringene i 2004 og 2007 for både UiB og UiO, til tross for at BORA ble innført i 2005, mens UiO ble innført i 2002.



Figur 27 - årsfordeling av snitt antall siteringer for IA/ikke-IA artikler for UiB og UiO.

Kurvene i Figur 27 viser årsfordelt snitt antall siteringer pr artikkel og forfattere for henholdsvis UiB og UiO. Jeg gjør oppmerksom på at grafen for UiB kun inneholder tallmaterieell fra emnekategoriene MAT og MED, mens UiO's graf inneholder materiale fra

kategoriene MAT og SAMF. UiB hadde ikke match mellom ISI og NORA for emnekategori SAMF, mens UiO ikke hadde match for emnekategori MED. Dette er tatt hensyn til i beregningene, og derfor heller ikke med i totaltallene for Norge eller i grafene.



Figur 28 – årsfordeling av artikkelutvalget fra ISI fordelt pr publiseringsår for UiB og UiO.

Forklaringen kan likevel forsvares når man ser tidspunktet for retningsendringene i sammenheng med årsfordelingen av artikkelutvalget fra ISI, som er grunnlaget for snittberegningene – se Figur 28 (Figur 28 tilsvarer Figur 7, men er gjengitt her som ny figur for å bedre lesbarheten). Ser man på kurvene for UiB først, er det en markert økning av

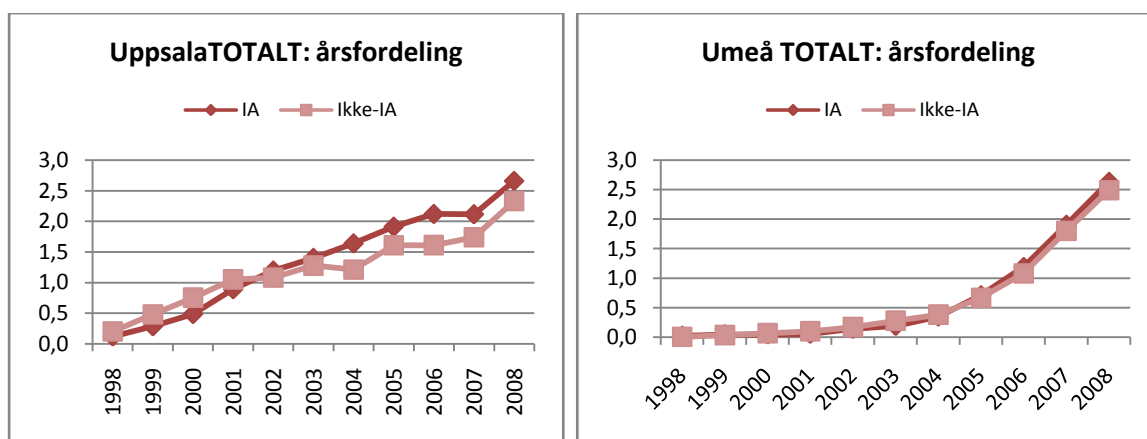
både IA og ikke-IA artikler med ”topper” både i 1999, i 2002 og i 2004 som kan gjenspeiles i økt antall siteringer et par år etter publisering når artiklene ”oppdages”. For UiB kan man altså ikke si noe om effekten av innføring av IA, som hadde oppstart først i 2005, før man eventuelt kan knytte det til retningsendringer i årene etter innføringen i 2005. Det kan være litt tidlig å antyde noe basert på mitt tallmateriale fra denne perioden, men sammenlignes snitt antall siteringer for IA artikler og ikke-IA artikler for UiB i 2007 har IA artiklene et snitt som ligger 36,9 % høyere enn ikke-IA artikler, hvor snittet var henholdsvis 2.9 for IA artikler og 2.1 for ikke-IA artikler. I 2008 øker forskjellen ved at IA artiklene har et snitt som ligger dobbelt så høyt i forhold til ikke-IA artiklene (98,5 %). Snittet i 2008 var henholdsvis 5,4 for IA artikler og 2,7 for ikke-IA artikler.

For UiOs vedkommende er tilsiget av artikler i ISI mer jevn – se UiOs graf i Figur 27, noe som gjenspeiles i en jevnere årsfordelt siteringskurve. Retningsendringen med en økning av antall siteringer både for IA og ikke-IA artikler i perioden 2004-2006 kan sannsynligvis tillegges økningen i antall artikler i 2003. Det samme kan antageligvis sies om toppen av antall artikler i 2007, som gir økte siteringer i 2008 og forhåpentligvis fremover, men det får tiden vise. Her er det også interessant å følge utviklingen i snitt antall siteringer etter innføring av IA, som for UiO var 2002. Gjør man her den samme sammenligningen av snitt antall siteringer for IA artikler og ikke-IA artikler hvert av årene fra et par år etter innføring av IA viser tallene også for UiO at det har vært en større økning i snitt antall sitering for IA artikler

enn for ikke-IA artikler. I 2006 var IA artikler 37,7 % mindre sitert enn ikke-IA artikler, hvor snittet var på henholdsvis 0,8 og 1,2 for IA/ikke-IA artikler, men i 2007 hadde snittet på IA artikler tatt igjen snittet for ikke-IA artikler, slik at snittet lå likt med et snitt på 1.0 for begge gruppene. I 2008 økte forskjellen ved at IA artiklene hadde et snitt som lå 63,9 % høyere i forhold til ikke-IA artiklene. Snittet i 2008 var henholdsvis 3,0 for IA artikler og 1,8 for ikke-IA artikler.

Økningen i siteringer ser altså ut til å følge ett par år etter økningen i antall artikler, og økningen i siteringer er høyere for IA artikler enn for ikke-IA artikler. Hvis denne tendensen vedvarer de neste årene mener jeg at man kan antyde at innføringen av IA har en positiv effekt, men fordi datagrunnlaget er lite, og veldig ”ferskt”, kan jeg ikke tolke det til annet enn en svak tendens.

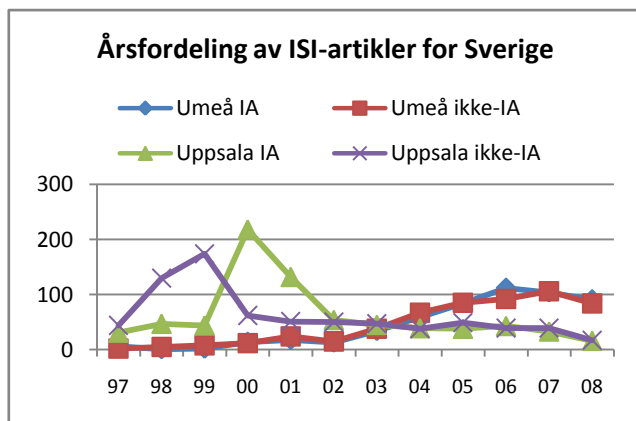
Grafen som viser årsfordelt snitt antall siteringer for Sverige forteller en litt annen historie – illustrert i Figur 25 på side 56. Her kan man også observere en liten retningsendring i 2003 for ikke-IA artikler med en utflating før ny stigning. For IA artikler oppstod ikke en tilsvarende endring, men en jevn stigning, og dermed har antall siteringer økt og gått forbi, dvs. at IA artikler har et høyere snitt antall siteringer pr artikkel enn ikke-IA artikler etter 2003.



Figur 29 - årsfordeling av snitt antall siteringer for IA/ikke-IA artikler for Uppsala og Umeå

Går man ned et nivå og ser på institusjonsnivå, se Figur 29, ser man at det er Uppsala som er årsak til endringen på landsnivå i 2003. Umeå har en jevnere kurve for både IA og ikke-IA artikler, og det ser ikke ut til at innføring av IA har hatt nevneverdig innflytelse på snitt antall siteringer, utover at snittet for IA artikler etter 2006 har økt og ligger litt foran snittet for ikke-IA artikler. For Uppsala er det en mer tydelig retningsendring for ikke-IA artikler i 2003, ikke mye, men nok til at man kan spørre om innføring av IA i 2002 likevel har hatt innvirkning på

antall siteringer. Det som kan tale for at innføring av IA har hatt en positiv effekt er at kurven for IA artikler steg før innføring og stiger etter innføring, dvs. ingen retningsendring, mens ikke-IA artikler ”lider” med en liten nedgang før også kurven her stiger noe mer ujevnt.



**Figur 30 - årsfordeling av artikkelutvalget fra ISI fordelt på publiseringsår for Uppsala og Umeå.**

Noe av forklaringen kan ligge i grafen for årsfordeling av artikkelutvalget fra ISI for Uppsala og Umeå, se Figur 30 (Figur 30 tilsvarer Figur 8Figur 7, men er gjengitt her som ny figur for å bedre lesbarheten). Sammenligner man kurvene i denne grafen med kurvene i grafen for årsfordelt snitt antall siteringer (Figur 29), ser man at Umeå har hatt et like jevnt tilsig av artikler som økningen av siteringer for både IA og ikke-IA artikler.

For Uppsala ser vi en økning i tilsiget av artikler i perioden 1998-1999 for ikke-IA artikler og tilsvarende for IA artikler i 2000. Normalt tar det et par år før en artikkel blir sitert. Det som er interessant er at kurven for antall siteringer for ikke-IA artikler for Uppsala ikke stiger nevneverdig i 2000-2002, eller i tilsvarende grad som økningen i antall artikler kanskje burde tilsi. Man kan også si at kurven for antall siteringer for IA artikler kanskje burde øke mer enn hva som er tilfelle i 2001-2003 for å demme opp for økningen i antall artikler i 2002. Det som derfor er viktigere her er at kurven er jevnt stigende i hele måleperioden og ikke får de samme negative retningsendring eller utflatingene som kurven for ikke-IA artikler i 2004 og 2006.

For Sverige ser også økningen i siteringer ut til å følge ett par år etter økningen i antall artikler på samme måte som for Norge, selv om det ikke er like tydelig. Økningen i snitt antall siteringer er også blitt høyere for IA artikler enn for ikke-IA artikler et par år etter innføring av IA. Resultatene er ikke like tydelige som for Norge, og jeg kan ikke tolke det til annet enn at det gir indikasjon på en svak tendens.

### 5.3 Problemstillinger og påvirkningsfaktorer

Hovedutfordringene i denne undersøkelsen har uten tvil vært datamaterialet. Jeg har allerede beskrevet analyseutvalget og svakhetene ved dette i kapittel 4.1 Prosjektets materiale. Jeg diskuterer derfor ikke utvalget mer her, men vil komme tilbake til både det og konsekvensene under presentasjonen av erfaringer med valg av metode og i oppsummeringen.

Det finnes flere andre faktorer som kan påvirke antall siteringer og dermed svekke min hypotese. Jeg trekker frem noen av punktene jeg mener er mest interessante, men listen er ikke fullstendig.

1. Håndtering av dubletter i IA og søketjenester.
2. Ulike synspunkter på IA og parallellpublisering.
3. Valg av siteringsdatabase og kriterier for hvilke tidsskrift som indekseres.
4. Trender som påvirker valget av tidsskrifter og artikler til indeksering.
5. Faktorer som er med å påvirke den lave andelen av artikler med match mellom ISI og NORA/DiVA.
6. Forskernes egenutvikling og mulighet for påvirkning av egen anerkjennelse.
7. IA i konkurranse med andre formidlingskanaler - nyhetens interesse.

Punktene presenteres og utdypes mer detaljert i fortsettelsen.

1. Jeg nevnte ovenfor at jeg ikke skulle diskutere datamaterialet i dette kapittelet. Jeg ønsker imidlertid å diskutere et spesielt område; problemene knyttet til dubletter i de institusjonelle arkivene. Siden NORA høster og DiVA samler fra flere institusjonsarkiver, og disse systemene ikke driver dublettkontroll, resulterer det i at en og samme artikkel kan forekomme flere ganger. Det er for så vidt riktig når en artikkel har flere forfattere, og disse er ansatt ved ulike institusjoner. Da vil artikkelen bli registrert i hvert av de institusjonelle arkivene i det antall arkiver som forfatterne er tilknyttet institusjoner. Et eksempel er artikkelen ”Curriculum factors influencing knowledge of communication skills among medical students” fra 2007 som er forfattet av 9 forfattere. 2 av forfatterne var på tidspunktet for registrering ansatt ved universitetet i Tromsø, og de andre ved universitetet i Bergen. Artikkelen er registrert både hos UiB i BORA og hos UiT i MUNIN, og forekommer derfor som to innførsler i NORA – én pr institusjon. Fenomenet gir altså dubletter ved innhøsting, og antall ganger en artikkel forekommer øker for hvert

nivå av innhøsting. NORA høster fra de norske IAene og risikerer dubletter fra de ulike norske institusjonene. Internasjonale høstere som OAIster og Bielefeld Academic Search Engine (BASE) høster både fra NORA og fra IAer ved universitetene, og risikerer derfor dubletter fra flere nivåer under innhøsting fra de ulike IAene. For mitt datamateriale er ikke dette et stort problem. Kun 2 artikler av totalt 377 i NORA var registrert under 2 institusjoner. Et større problem er de dublettene som er forårsaket av dobbeltregistrering hos det enkelte arkiv, enten fordi samme artikkel registreres tilhørende flere avdelinger eller rett og slett på grunn av feilregistrering. Igjen ikke et stort problem i det norske materialet. Kun 9 av artiklene som var dobbeltregistrert innen samme arkiv var rene dubletter, dvs. kun 2,4 % av artiklene i NORA, I DiVA er dette et mer utstrakt fenomen. 12,8 % av artiklene (578 artikler av totalt 4500 artikler i utvalget) var dobbeltregistrert. Jeg har ikke tatt hensyn til dette i analysegrunnlaget, men ser at det er en faktor som kan ha hatt en mulig påvirkning på resultatet for Sverige.

2. Hvem egenarkiverer/parallellpubliserer i institusjonelle arkiver i dag? Det er en interessant problemstilling sett i lys av mengden artikler som er innhøstet til NORA, dvs. de artiklene som utgjør min populasjon. Dersom IA møtes med så mye skepsis blant forskerne som det gis inntrykk av gjennom kritikken og forbeholdenheten i avisartikler (Bjørkeng 2009), og det faktum at lite blir parallellpublisert i forhold til den faktiske produksjonen, er det interessant å vite hvem som arkiverer. Er det de ”nye” og ”ferske” forskerne som enda ikke er så kjente, og som ikke har fått et ”navn”, eller er det de ”gamle” og ”etablerte” forskerne som har publisert mer, og derigjennom hatt en større sjanse for å bli sitert? Antall siteringer knyttet til IA artikler kan være påvirket av hvilke forskergrupper som bruker IA bevisst. Antall siteringer knyttet til IA artikler kan bli forskjellig avhengig av om det er de yngste forskerne med liten produksjon som velger å parallellpublisere i IA eller om det er de eldste med lang fartstid. Det er en viss sannsynlighet for at de eldste har et større antall siteringer enn de yngste, selv om jeg ikke har gått inn i problematikken i tilknytning til mitt datautvalg. Det kan være en interessant undersøkelse i seg selv. Det jeg har registrert er at det på tidspunktet for utplukk var registrert kun 2 fagfelleverderte artikler i NORA i emnekategori medisiniske fag ved UiO, og det gjenspeiler på ingen måte artikkelproduksjonen i dette fagmiljøet. Hvordan hadde tallene sett ut om alle artiklene publisert fra dette miljøet også hadde vært parallellpublisert i DUO hos UiO?

3. Jeg har valgt å hente siteringsdata fra ISI Web of Knowledge og databasene Science Citation Index Expanded, Social Sciences Citation Index og Arts & Humanities Citation Index. Ved å bruke denne ene tjenesten er jeg prisgitt hvilke kriterier Thomson Reuters har satt for utplukk av tidsskrifter til indeksering innen de ulike vitenskapsdisiplinene. ISI er stor, og tjenesten favner et bredt spekter av tidsskrifter over mange år, men det er en internasjonal tjeneste for et engelskspråklig publikum, så nordiske tidsskrifter hvor artiklene skrives på et av de nordiske språkene er lite representert. Det resulterer i at antallet artikler med match mellom NORA/DiVA og ISI også er så lite som det faktisk viser seg å være. Igjen kan tallene ha vært annerledes om jeg hadde valgt andre emnekategorier, eller om jeg hadde brukt andre eller flere forfattere enn de jeg har brukt. Jeg kunne også ha brukt andre mer fagrelaterte arkiver enn rene institusjonsarkiv, som for eksempel Helsebibliotekets forskningsarkiv, HeRa. Uansett, så lenge man bruker datamateriale med utgangspunkt i Norge vil det norske "tellekant-systemet"<sup>13</sup> ha en viss innvirkning. I Norge får forskerne og institusjonene "betalt" gjennom såkalte publikasjonspoeng, dvs. at man har en nasjonal pott til utbetaling knyttet til hvor mange publikasjonspoeng en institusjon oppnår. Poengene gis med bakgrunn i hvor en publikasjon blir publisert, type publikasjon og hvor stor andel av publikasjonen institusjonens forskere står bak. Det finnes oversikter over de tidsskriftene som gir poeng på to ulike nivåer, hvor nivå 2 gir mest uttelling, men hvor det også er vanskeligst å få publisert. Her finner man de internasjonale, prestisjetunge tidsskriftene, som igjen er de som i størst grad indekseres i ISI. I en ny studie hadde det vært interessant å få undersøkt hvor mange av nivå 1 og nivå 2 tidsskriftene som faktisk blir indeksert av ISI, og om det finnes andre siteringstjenester som også indekserer de samme tidsskriftene.
4. Jeg beskrev over problemstillingen knyttet til hvor liten innflytelse man har på valg av tidsskrifter til indeksering hos ISI Web of Knowledge. Selv om Thomson Reuter overvåker markedet og henter inn tidsskrifter med stor anseelse og fjerner de som har mistet den, vil alltid "ting i tiden", dvs. tematikk som er populær i media for øvrig, gis oppmerksomhet. De siste årene har for eksempel temaene klima, miljø, genteknologi og finanskriser vært høyt oppe på nyhetstoppen, og ISI vil selvsagt vurdere de tidsskriftene som fokuserer på disse temaene. Forskning som beskjeftiger seg innen disse fagområdene

---

<sup>13</sup> Tellekantsystemet er en finansieringsmodell for forskning ved universiteter og høyskoler. Systemet belønner artikler som publiseres i tidsskrifter tilhørende én av to definerte grupper av tidsskrifter. Artikler som trykkes i prestisjetunge tidsskrifter (nivå 2) får mest uttelling. Systemet står for 2 prosent av forskningsfinansieringen på landsbasis.

får større fokus, og artikler som tar opp tematikken vil ha en høyere leserfrekvens og sannsynligheten for å bli sitert er større. Dette er en faktor som kan påvirke antall siteringer og et datagrunnlag som baserer seg på uttrekk fra en bestemt tidsperiode, dvs. om valgt periode omfatter årene som tar opp i seg slik ”ting i tiden”-trender. Finanskrisen er en forholdsvis fersk nyhet, så artikler tilknyttet dette temaet er nok fortsatt ikke registrert i IA og ikke oppnådd å få siteringsdata indeksert i ISI. Tematikken rundt miljø og klima derimot har pågått noen år, så her kan artikler ha blitt parallellpublisert og siteringsdata indeksert, selv om siteringene gjelder for et lite antall år, og dermed er få i antall.

5. Et annet interessant spørsmål som har dukket opp i forbindelse med analysearbeidet er hvorfor det er så mange av artiklene i ISI som ikke er parallellpublisert? Jeg fant 2111 artikler i ISI for de 118 forfatterne i forfatterutvalget mitt. I NORA var disse registrert med 321 artikler når jeg fordelte artiklene pr forfatter. Det betyr at antall artikler i NORA relatert til disse forfatterne utgjorde kun 15,2 % av artiklene indeksert ISI, eller sagt på en annen måte, NORA ”mangler” 84,8 % av artiklene i ISI. Hva er årsaken? Har man ikke kommet i gang med registrering, eller velger man å ikke registrere bakover i tid? Kan det bety at noen av disse artiklene venter på å bli lagt inn etter dato for utløpt embargo<sup>14</sup>? Hvem foretar valg for hva som skal/ikke skal parallellpubliseres, eller er det rett og slett mange av artiklene i ISI som allerede ligger åpne fordi de er publisert i OA tidsskrifter? Antallet artikler i NORA, eller mangel på artikler, har hatt konsekvenser for datagrunnlaget og analysene mine, men jeg stiller bare spørsmålene her. Omfanget av oppgaven har ikke tillatt meg å gå videre inn i problematikken, men igjen en interessant problemstilling som kan belyses med nye undersøkelser.
6. En annen faktor som kan påvirke antall siteringer er forskernes egenutvikling og deres mulighet til selv å påvirke egen anerkjennelse. Jeg nevnte tidligere de ”nye” og ”ferske”, som ikke har rukket å publisere så mange artikler enda, og som dermed ikke har rukket å bli kjent eller fått et ”navn”, og de ”gamle” og ”etablerte” som har publisert mer, og derigjennom hatt en større sjanse for både å bli sitert og fått et ”navn”. Forskertyper er det mange av, og over tid har det vært og det vil sikkert også i fremtiden være store forskjeller på publiseringshyppighet. Et positivt resultat av innføring av ”tellekant-systemet” her i

---

<sup>14</sup> **Embargo** er en blokada, dvs. en artikkel er underlagt en embargo/blokade, eller et **forbud** mot parallellpublisering inntil en gitt dato. Ofte er embargotiden gitt for en tidsperiode på for eksempel 3, 6 eller 12 måneder etter publisering.



Norge er en økning i antall publiserte fagfelleverderte dokumenter siden registreringen av publisering startet i 2004. I en artikkel i Aftenposten (Lund 2009) kommer det frem at den vitenskaplige produksjonen i universitets- og høyskolesektoren har økt med 12,5 %, og at publiseringen i nivå-2 tidsskrifter (vurdert som de mest prestisjetunge), har økt med hele 50 % siden 2005. Når jeg ser på tallene som benyttes i artikkelen har antall tidsskriftsartikler økt med 6,6 % fra 2007 til 2008, og 21,2 % fra registreringen startet i 2004. Økt produksjon vil på sikt forhåpentligvis gi økt antall siteringer, og dermed økt anerkjennelse for norsk forskning. Om alle artiklene det refereres til også er parallellpublisert har jeg ikke sjekket. Det jeg vet er at NORA økte antall dokumenter totalt fra 19698 til 22067 på 8 mnd fra 06.10.08 til 02.05.09. Det er en økning på 10,7 %, så fortsatt er det en jobb å gjøre med strategien for parallellpublisering.

7. Det siste jeg skal nevne er muligheten for at innføring av IA i seg selv har hatt en påvirkning på antall siteringer. IA er en ny formidlingskanal som markedsføres av institusjonen og institusjonens bibliotek, som ofte forvalter IAene, og dermed i en periode er og har vært av nyhetens interesse. Kan det i seg selv gi et hopp i antall siteringer? Dersom så er tilfelle, er hensikten med innføringen av IA oppnådd. Synliggjøring er stikkordet for arkivene, - synliggjøring av den enkelte forskers, faggruppes og institusjonens vitenskaplige produksjon. Om innføringen i seg selv har hatt påvirkning på antall siteringer, så er det en faktor som er inkludert i tallmaterialet mitt og er en del av undersøkelsen. I Norge er IA et forholdsvis nytt fenomen, og det er betimelig å spørre hva som skjer når nyhetens interesse har lagt seg og hverdagen igjen er der? BORA startet opp i 2005. DUO har eksistert siden 2002, og det samme har DiVA hos universitetene i Uppsala og Umeå. Resultatene fra analysene viser en trend med økende antall siteringer for alle institusjonene uavhengig av starttidspunkt, men det er klart at en ny undersøkelse om noen år med større datagrunnlag, og når IA som kanal har eksistert lengre, vil gi bedre svar på om innføring kun ga et blaff i økningen av antall siteringer.

#### **5.4 Erfaringer med bruk av ”ny” metode**

For å gjennomføre mine studier i innføring og bruk av IAer valgte jeg å utvikle og bruke en ”ny” metode. I stedet for den ”vanlige” metoden med utgangspunkt i et utvalg med vitenskaplige tidsskrifter, hvor man teller sitering pr artikkel og sjekker om disse artiklene er parallellpublisert i IA eller ikke, har jeg valgt å ta utgangspunkt i det som allerede er

parallellpublisert. Jeg teller antall siteringer for disse artiklene og sammenligner antall siteringer for artikler parallellpublisert med siteringer for artikler av samme forfatter som ikke er parallellpublisert. Begge metodene gir mulighet til å undersøke ”research impact”, eller anerkjennelse, sett i lys av IA som kanal for parallellpublisering. Her er det altså veien til målet som har vært forskjellig. Jeg mener at det kan oppfattes som støy i datagrunnlaget å sammenligne siteringer på artikler i tidsskrifter publisert av ulike forfattere fra mange ulike institusjoner. Etter all sannsynlighet vil tidsskriftenes forleggere, artikkelforfattere og forfatternes institusjoner ha veldig ulik oppfatning av og holdning til bruk av Open Access, men dette er kunnskap vi ikke har. Når jeg konsentrerer meg om den enkelte forsker og hans/hennes fagfelt innen en institusjon vil datagrunnlaget være mer homogent. Den ”nye” metoden sammenligner forfattere, faggrupper og institusjoner med seg selv, dvs. at siteringer for artikler som er parallellpublisert i IA og artikler som ikke er det for samme forfatter sammenlignes. Dette kan skaleres slik at man sammenligner siteringer for flere forfattere / faggrupper innen samme institusjon, eller for flere forfattere / faggrupper fra flere institusjoner. Metoden gjør det enkelt å sjekke forfatternes og institusjonenes oppfatning av og holdning til Open Access og parallellpublisering, og det er enkelt å velge hvilket nivå som er hensiktsmessig for en undersøkelse.

Siden metoden er ny har det vært et viktig element i gjennomføringen å teste selve metoden, dvs. diskutere om denne måten å gjøre det på egner seg til å belyse de problemstillingene jeg tar opp i tilknytning til IA og ”impact”/ anerkjennelse.

Den første og kanskje viktigste erfaringen er at metoden er svært arbeidskrevende dersom det ikke benyttes verktøy/programmer som forenkler uttak fra databasene og gjennomføring av analysene, som for eksempel SPSS. Jeg var nok ikke klar over omfanget, og fikk ikke testet dette godt nok før det var for sent å legge om. Derfor er alle uttak og beregninger gjort manuelt, og det eneste verktøyet jeg har benyttet er Microsoft Office Excel. I tillegg til omfanget av tilrettelegging av analysegrunnlaget har det tatt mye tid å kvalitetssikre og foreta beregninger i forbindelse med vurdering av analysegrunnlaget. Dette ville vært lettere å håndtere i et analyseverktøy bedre tilpasset formålet. Fordelen har vært at jeg har hatt fullstendig ”hands on”, men det veier ikke opp for alle timene med regnearkene.

En annen viktig erfaring, som kanskje ikke har noe med metoden å gjøre, men med min uerfarenhet med statistisk datamateriale, var at jeg først etter at jeg hadde foretatt analysene

oppdaget at datamateriale fra universitetet i Tromsø (UiT) og Høgskolen i Telemark (HiT) ikke kunne benyttes. Oppstartstidspunktet for de institusjonelle arkivene MUNIN hos UiT og TEORA hos HiT var så sent i tidsintervallet for undersøkelsen at analyser av antall siteringer før og etter innføring av IA ikke hadde noen hensikt. MUNIN hadde oppstart i 2007, og TEORA i 2008. Se Tabell 1 – Oversikt over nasjonale og lokale IAer og deres offisielle oppstartstidspunkt. Konsekvensen var at alle analysene måtte gjennomføres på nytt med et mindre og endret analysegrunnlag. Her hjalp heldigvis regnearkmodellene meg, som jeg hadde utviklet under bearbeiding av datamaterialet, slik at operasjonen med å fjerne data fra de to institusjonene var forholdsvis enkel.

Jeg har vært opptatt av forskerne/forfatterne og institusjonene. I presentasjonen av de ulike analysene har jeg kun brukt de akkumulerte nivåene land, institusjon og emne kategorier. Alle beregninger starter på individnivå, for så å bli summert opp på de øvrige nivåene, men det har ikke vært hensiktsmessig å trekke frem individnivået i presentasjonen. Utplukket av forfatternavn til utvalget har vært tilfeldig. Jeg satte noen kriterier for utplukket, slik som tilknytning til institusjon og emne kategori, men andre kriterier kan fungere like bra. Jeg brukte forfatternavn basert på parallellpubliserte artikler i IA, men den beste løsningen, og den mest kvalifiserte i henhold til statistiske teknikker, er å bruke alle ansatte, eller en delmengde av ansatte i forskningsstillinger ved de institusjonene som velges. Også utplukket av institusjoner til utvalget har vært tilfeldig. Jeg satte kriteriet til at institusjonen måtte ha en viss mengde artikler innhøstet til NORA for å gi meg en viss mengde data å ta utgangspunkt i. Emne kategoriene ble valgt ut fra det samme kriteriet. For å få et utvalg med bedre spredning og statistisk representativitet enn det jeg oppnådde i denne undersøkelsen, bør kriteriene for plukk av datamaterialet gjennom en grundig vurdering i forkant av en ny undersøkelse for å sikre kvaliteten i analyseutvalget. Min datamengde ble ikke stor nok, og jeg burde benyttet flere forfattere for å få et større utvalg av artikler og siteringsdata.

Et lite paradoks til slutt. Når man jobber med Open Access, dvs. med åpen og fri tilgang til data, er det spesielt å jobbe med siteringsdatabasene i ISI Web of Knowledge, som ikke er åpne og frie. Tjenestene i ISI er ikke gratis, og nås kun gjennom en abonnementsordning hvor det betales i forhold til hvilket behov man har. Høgskolen i Oslo abonnerer på ISI, men tilgangen er knyttet til søk og ikke uttrekk av større mengder data. Drømmen hadde vært å få sendt en fil med alle forfatternavnene, og i retur fått en fil med alle artiklene for disse forfatterne med siteringsdata som automatisk kunne lastes inn i SPSS eller et annet analyse-

verktøy. Situasjonen var ikke slik i mitt tilfelle. Alt materialet fra ISI ble hentet inn via kommaseparerte datastrenger som jeg fikk via elektronisk post basert på enkeltsøk på forfatternavn sjekket mot ISI's autoritetsregister. Jeg undersøkte muligheten for å finne siteringsdata i andre databasetjenester, men ingen har de mengdene som ISI kan tilby. Erfaringsmessig vil jeg nok anbefale for eventuelle senere undersøkelser å følge med på utviklingen av tilgjengelige siteringsbaser, i tillegg til å se på muligheten for "drømmeuttrekket" fra ISI beskrevet over. Hvilken siteringsdatabase man velger og hvordan man får tilgang vil være den samme problemstillingen for både den "vanlige" og den "nye" metoden. Om utgangspunktet er tidsskrifter eller forfatternavn vil utfordringene være de samme.

## 6. Oppsummering og konklusjon

"Institusjonelt arkiv – en kanal med impact?" Hovedformålet mitt har vært å undersøke om innføringen av IA har hatt en positiv effekt på gjenfinning og bruk av institusjonens forskningsproduksjon med utgangspunkt i norske forhold. Jeg har hatt to hensikter med gjennomføringen av undersøkelsen. For det første ønsket jeg å teste en hypotese tilknyttet IA som formidlingskanal og IAens mulige effekt på antall siteringer. For det andre ønsket jeg å teste en ny metode tilknyttet bruk av IA og måling av "impact"/anerkjennelse.

Min hypotese har vært at vitenskaplige artikler produsert ved norske universitet og høyskoler får økt "citation impact" – flere siteringer, dersom artiklene er åpne og fritt tilgjengelige gjennom et IA, enn om de ikke er det. Jeg stilte spørsmål om i hvilken grad antall siteringer for artikler som er parallellpublisert i IA (IA artikler) vil øke sammenlignet med artikler som ikke er parallellpublisert (ikke-IA artikler), dvs. kun publisert i tidsskrifter som er den tradisjonelle kanalen for publisering av vitenskaplige fagfelleverderte artikler. Jeg stilte også spørsmål om i hvilken grad IA artikler blir sitert raskere sammenlignet med ikke-IA artikler.

Utgangspunktet i metoden som er brukt er forskere og deres vitenskaplige produksjon. Metoden sammenligner forfattere, faggrupper og institusjoner med seg selv, dvs. at siteringer for artikler som er parallellpublisert i IA og artikler som ikke er parallellpublisert fra samme forfatter sammenlignes. Dette kan skaleres slik at man kan sammenligne siteringer fra flere forfattere / faggrupper innen samme institusjon, eller fra flere forfattere / faggrupper fra flere

institusjoner. Sammenligningen gjøres ved at siteringer for et antall vitenskapelige artikler for et utvalg av forskere telles, og siteringshastighet analyseres og sammenlignes for artikler som er parallellpublisert i IA med artikler som kun er publisert i tradisjonelle kanaler. Metoden er skalerbar og antall siteringer aggregeres opp til det nivået som er hensiktsmessig for den undersøkelsen som skal gjennomføres. I min undersøkelse er det aggregert og foretatt sammenligninger på emnekatgorinivå, institusjonsnivå og landnivå.

Datagrunnlaget i undersøkelsen baserer seg på et utplukk av fagfellevurderte tidsskriftsartikler fra NORA/DiVA, hvor artikkelforfatterne utgjør grunnlaget i et forfatterutvalg.

Forfatterutvalget ble så brukt som utgangspunkt for videre søk i ISI og utplukk av alle artiklene med siteringsdata for disse forfatterne. Disse artiklene utgjør artikkelutvalget i undersøkelsen. Siteringene ble ordnet, summert og aggregert for å kunne analysere snitt antall siteringer fra forskjellige vinklinger. I alle sammenligninger operer jeg med to grupper. Den ene av grupperingene kaller jeg ”IA artikler”, dvs. artikler som både finnes i ISI og NORA/DiVA (artikler med match), og den andre grupperingen for ”ikke-IA artikler”, dvs. de artiklene jeg har plukket fra ISI, som kun finnes i ISI og som ikke er parallellpublisert.

Jeg stilte to spørsmål for å kunne teste og måle hypotesen. Det første spørsmålet; om i hvilken grad antall siteringer øker for IA artikler sammenlignet med ikke-IA artikler ble forsøkt besvart gjennom to ulike analyser. Den første analysen så på eventuelle endringer i antall siteringer før og etter tidspunktet for innføring av IA uavhengig av om artiklene var parallellpublisert eller ikke, mens den andre så på eventuelle forskjeller i antall siteringer for IA/ikke-IA artikler.

1.1. Resultatet i den første analysen viste, sett alle artiklene under ett, dvs. både IA og ikke-IA artikler, at det er en tydelig tendens med økning av snitt antall siteringer i perioden etter innføring av IA. Til tross for at fordelingen på antall artikler i det norske utvalget var på henholdsvis 53,4 % og 46,5 % for periodene før og etter innføring av IA, slik at en økning i antall siteringer vil være forventet basert på erfaringen med at de første siteringene som oftest inntreffer et par år etter publisering, så steg snitt antall siteringer med hele 151,5 %. Selv når artiklene grupperes i kategoriene før og etter, dvs. artikler publisert før innføring av IA og artikler publisert etter innføring, og man ser på fordelingen av antall siteringer og sammenligner snitt antall siteringer de første årene etter publisering for disse kategoriene så kommer tendensen tydelig frem. For

Norges vedkommende økte antall siteringer etter innføring av IA for før-gruppen med 67,8 %, til tross for at periodene det telles siteringer for er på 6,5 år før innføring av IA og på 5,5 år etter innføring. Ser man på antall siteringer for gruppen etter, dvs. artikler publisert etter innføring av IA, så er det her en reduksjon i antall siteringer med 16,3 %. Reduksjonen kan være knyttet til at både artikkelfordelingen og antall år det telles siteringer for er skjev mellom de to måleperiodene. Sammenligner vi dette med Sverige var det også der en vesentlig økning i antall siteringer etter innføring av IA. Økningen totalt i snitt antall siteringer var på hele 456,3 % når man ser alle artiklene under ett, mens økningen var på 204,7 % for gruppen av artikler publisert før innføring i IA. Økningen i antall siteringer etter innføring av IA på 152,2 % er fortsatt vesentlig, men er nok påvirket av at antall artikler i det svenske utvalget hadde en annen fordeling enn det norske, med henholdsvis 40,0 % og 60,0 % for periodene før og etter.

1.2. Denne tendensen bekreftes når jeg sammenligner snitt antall siteringer for IA artikler med ikke-IA artikler. Når jeg foretar sammenligningen basert både på snitt antall siteringer for alle årene fra artiklenes publiseringsår til og med 2008, og på snitt antall siteringer for de tre første årene etter artiklenes publisering, viser tallene en økning i antall siteringer for IA artikler i forhold til ikke-IA artikler. Dette gjelder for både Norge og Sverige. For Norge er økningen på 94,3 % basert på beregningen av snitt antall siteringer fra artiklenes publiseringsår frem til 2008, mens den er på 87,0 % basert på beregningen av snitt antall siteringer de 3 første årene etter publisering. Tilsvarende resultater fra Sverige er ikke like høye, men også der er økning i antall siteringer for IA artikler høyere enn antall siteringer for ikke-IA artikler. Økningen er på 8,6 % for snittet på siteringer for perioden publiseringsår-08, mens økningen er på 7,1 % for snittet på antall siteringer for de 3 første leveårene.

Som svar på det første spørsmålet kan jeg på bakgrunn av resultatene av disse analysene konkludere med at tendensen til økninger i antall siteringer for parallellpubliserte artikler er vesentlig. Økningen kan ikke alene forklares med fordelingen av artikler i forhold til tidspunkt for før og etter innføring av IA eller lengden på siteringsperiodene. Jeg mener at tallene gir god indikasjon på at det er sannsynlighet for et høyere i antall siteringer for artikler parallellpublisert i IA i forhold til artikler som ikke er parallellpublisert.

Det andre spørsmålet jeg stilte for å kunne teste og måle hypotesen; om IA artikler blir raskere sitert sammenlignet med ikke-IA artikler, ble forsøkt besvart gjennom tre ulike analyser hvor jeg så på siteringshastigheten fra ulike vinklinger.

2.1 Først sjekket jeg hvor stor prosent av både IA og ikke-IA artiklene som ble sitert hvert år de 3 første årene etter publisering, og sammenlignet de to grupperingene.

Resultatene viser at sannsynligheten for å bli sitert er høyere hvert av de 3 første årene etter publisering dersom artikkelen er parallellpublisert enn om den ikke er det. Dette gjelder for både Norge og Sverige. For Norges vedkommende ble resultatet en økning i siteringshastigheten for IA artikler sammenlignet med ikke-IA artikler hvor andelen av siteringer for IA artiklene lå 48,6 % høyere første år, 18,9 % høyere andre år og 30,5 % høyere tredje år. Tallene fra Sverige viser samme tendens. Økningen i siteringshastighet for IA artikler i forhold til ikke-IA artikler er noe lavere enn i Norge, men økningen er fortsatt på 18,8 % første år, 9,5 % andre år og 5,4 % tredje år.

2.2 En annen vinkling var å se på om innføring av IA hadde hatt effekt på antall siteringer for ”gamle” artikler, dvs. artikler publisert før innføring av IA. Sagt på en annen måte; om artiklene fikk en ”ny giv”, eller oppsving i antall siteringer etter innføring av IA. Jeg målte dette ved å sammenligne antall siteringer i perioden før og etter innføring av IA for et utvalg av artikler publisert minimum 3 år før innføring av IA. Resultatene viser en svak positiv utviklingstendens, og viser at det skjer et hopp eller økning i antall snitt siteringer et par år etter tidspunktet for innføring av IA. Jeg kan likevel ikke generalisere funnene, fordi utvalget jeg baserer analysen på er sterkt redusert i forhold til utvalget for undersøkelsen som helhet. De økningene i antall siteringer jeg finner for IA artikler kan skyldes innføring av IA, men også andre faktorer kan ha hatt innvirkning. Både IA og Open Access som fenomen er nye, og det vil enda ta noe tid før dette er allment kjent. Tallmaterialet kan antyde at jeg har vært for tidlig ute med en ”ny giv”-analyse, og at dette bedre lar seg analysere noen år frem i tid.

2.3 Selv om jeg ikke kan generalisere resultatet fra ”ny giv”-analysen, så bekreftes trenden med en økning i antall siteringer for IA artikler i forhold til ikke-IA artikler også i den siste analysen. Her så jeg på utviklingen av snitt antall siteringer fordelt pr år fra publiseringsåret frem til og med 2008 for både IA og ikke-IA artikler i tidsintervallet 1998-2008. Det interessante er at tallmaterialet viser positive retningsendringer på

antall siteringer for begge kategorier artikler et par år etter innføring av IA. For Norges vedkommende ble DUO innført hos UiO i 2002 og BORA ble innført hos UiB i 2005. Både i 2004 og 2007 skjer det en økning i snitt antall siteringer for både IA og ikke-IA artikler, men økningen for IA artikler er høyere enn for ikke-IA artikler. Tallene for Sverige viser samme tendens, men forskjellen i antall siteringer mellom IA og ikke-IA artikler er mindre. Likevel ser man også der en sterkere økning for IA artikler sammenlignet med ikke-IA artikler. Retningsendringene skjer altså et par år etter innføring av IA, og man vil kunne anta at innføring av IA har hatt en positiv effekt. Tar jeg hensyn til fordelingen av antall artikler i det samme tidsintervallet er forskjellen så liten mellom IA og ikke-IA artikler at jeg skal være forsiktig med å vektlegge det for mye.

Som svar på det andre spørsmålet kan jeg på bakgrunn av resultatene fra disse analysene ikke konkluderer like sterkt. Tallene viser en positiv utviklingstendens for at IA artikler blir raskere sitert sammenlignet med ikke-IA artikler, men resultatene viser at forskjellen er så liten at det er vanskelig å generalisere. Her må eventuelle nye undersøkelser i fremtiden bekrefte eller avkrefte tendensen.

Veksten i antall siteringer for IA artikler i forhold til ikke-IA artikler kan selvsagt være påvirket av andre faktorer enn innføring av IA. Jeg har tatt utgangspunkt i alle artiklene som finnes i ISI for de forfatterne som er med i mitt forfatterutvalg. Jeg har ikke undersøkt i hvilke tidsskriftstyper artiklene er publisert i, og jeg har ikke sjekket publiseringspolicyen til forlagene. Det betyr at flere artikler allerede kan være publisert i OA tidsskrift, som gjennom en åpen og fri tilgang til artiklene gjør artiklene like tilgjengelige som parallellpublisering i et IA. Videre kan forfattere i mitt utvalg både være ”gamle”, etablerte og anerkjente forskere innen sitt felt, som både leses og siteres ofte uavhengig hvor artiklene er publisert, eller det kan være ”nye” forskere som parallellpubliserer alle sine artikler i håp om å komme raskere til ”markedet” for å bli lest og sitert. Flere mulige påvirkningsfaktorer kan nevnes, men felles for dem alle er at det er vanskelig å vekte disse slik at de tas høyde for i analysegrunnlaget. Hver faktor i sær kan være grunnlag for egne undersøkelser. Videre kan man stille spørsmål om det er andre elementer eller egenskaper som er felles for de artiklene som parallellpubliseres i IA sammenlignet med de artiklene som ikke parallellpubliseres. Har IA artikler unike særtrekk som man ikke finner hos ikke-IA artikler? Det er et interessant spørsmål som jeg dessverre ikke kan gi noen svar på, men ser det som interessant for fremtidig forskning.



Kausalitet har ikke vært mulig med mitt datamateriale, men det viser tendenser. Datautvalget i undersøkelsen er ikke statistisk representativt, slik at resultatene kun gir indikasjon på retning og viser utviklingstendenser. Resultatene gir indikasjon på at det er sannsynlighet for en økning i antall siteringer for artikler som er parallellpublisert i IA sammenlignet med artikler som ikke er parallellpublisert, og det er sannsynlighet for at IA artikler blir raskere sitert. Basert på de tendensene resultatene viser kan min hypotese beholdes, dvs. at vitenskaplige artikler får økt ”citation impact” – flere siteringer, dersom artiklene er åpne og fritt tilgjengelige gjennom IA enn om de ikke er det. Dersom man stiller også spørsmålet ”What’s in it for me?” blir svaret det samme. Trenden etter få år med institusjonelle arkiv i Norge tyder på at parallellpublisering kan gi bedre synlighet, flere siteringer, og dermed økt anerkjennelse.

Metoden som er brukt har gjennom de resultatene som er oppnådd vist at den egner seg. Det begrunner jeg med at resultatene jeg har fått på mine analyser viser de samme tendensene til økning i antall siteringer for artikler som er parallellpublisert som tidligere forskning har kommet frem til. Fortsatt kan metoden utvikles i forhold til valg og utplukk av datamateriale og bruk av analyseringsverktøy, men prinsippet med fokus på forskere, faggrupper og institusjoner passer godt for undersøkelser av publisering og tilgjengeliggjøring av vitenskaplig og faglig materiale i den norske universitets- og høyskolesektoren.

## 7. Litteraturliste

ABM-utvikling (2006). *Bibliotekreform 2014. Del 1: Strategier og tiltak*. I: ABM Skrift nr. 30. Lokalisert 18.09.08 på Verdensveven: [http://www.abm-utvikling.no/publisert/abm-skrift/bibliotekreform2014\\_del1.pdf](http://www.abm-utvikling.no/publisert/abm-skrift/bibliotekreform2014_del1.pdf)

ABM-utvikling (2009). *Norgesbiblioteket*. Lokalisert 18.05.09 på Verdensveven: <http://www.abm-utvikling.no/bibliotek/norgesbibliotek>

Antelman, K. (01.01.2004). *Do Open Access Articles Have a Greater Research Impact?* Lokalisert 15.10.08 på Verdensveven via OAIster: <http://www.citebase.org/abstract?id=64788591> eller <http://eprints.rclis.org/archive/00002309/>

Barwick, J. (20.01.07). *Building an institutional repository at Loughborough University: some experiences*. Loughborough, UK: Pilkington Library, Loughborough University. Article in *Program: electronic library and information systems*, Vol. 41, No. 2, pp. 113-123. Lokalisert 01.09.08 på OAIster: <http://hdl.handle.net/2134/2640>

Berlindeklarasjonen (20 - 22 Oct 2003). *Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*. Lokalisert 20.09.08 på Verdensveven: <http://oa.mpg.de/openaccess-berlin/berlindeclaration.html>

Bjørkeng, P.K. (27.02.09). *Forskere føler at de blir flådd. / Kjempebra fremstøt. / Mange er skeptiske*. I: Aftenposten (Morgenutg. Kulturdelen). Oslo, s. 6-7.

Brody, T. and Harnad, S. (Juni 2004). *Comparing the Impact of Open Access (OA) vs. Non-OA Articles in the Same Journals*. Article in *D-Lib Magazine*, Vol. 10, No. 6. Lokalisert 15.10.08 på Verdensveven via OAIster: <http://eprints.ecs.soton.ac.uk/10207/1/06harnad.html>

Buehler, M. A. & Boateng, A. (06.04.2005). *The evolving impact of institutional repositories on reference librarians*. Rochester, New York, USA: Wallace Library, Rochester Institute of Technology. Article in *Reference Services Review*, Vol. 33, No. 3, pp. 291-300. Lokalisert 01.09.08 på OAIster: <http://hdl.handle.net/1850/1369>

Cornell University Library (2008). *arXiv.org*. Lokalisert 18.10.08 på Verdensveven: <http://arxiv.org/>

DOAJ (2008). *Directory of Open Access Journals*. Lokalisert 01.09.08 på Verdensveven: <http://www.doaj.org/>

European Research Advisory Board (EURAB) 06.049 (2006). *Scientific Publication : Policy on Open Access*. Lokalisert 20.09.08 på Verdensveven: [http://ec.europa.eu/research/eurab/pdf/eurab\\_scipub\\_report\\_recomm\\_dec06\\_en.pdf](http://ec.europa.eu/research/eurab/pdf/eurab_scipub_report_recomm_dec06_en.pdf)

European Research Council (ERC) (2006). *ERC Scientific Council Statement on Open Access December 2006*. Lokalisert 20.09.08 på Verdensveven: <http://erc.europa.eu/pdf/open-access.pdf>

Fagerjord, K. (12.11.07). *Open Access – opprettelse av åpent institusjonelt arkiv i HiBu. Rapport fra arbeidsgruppe*. Drammen, Hønefoss og Kongsberg: HiBu. Lokalisert 01.09.08 på Verdensveven: [http://idtjeneste.nb.no/URN:NBN:no-bibsys\\_brage\\_5263](http://idtjeneste.nb.no/URN:NBN:no-bibsys_brage_5263)

Forskningsrådet (06.02.2009a). *Nye retningslinjer for digital publisering*. Lokalisert 28.04.09 på Verdensveven: <http://www.forskningsradet.no/no/Nyheter/Nye+retningslinjer+for+digital+publisering/1233558139716>

Forskningsrådet (2009b). *I front for forskningen. Strategi for Norges Forskningsråd 2009-2012*. Lokalisert 19.10.08 på Verdensveven: <http://www.forskningsradet.no/no/Strategi+og+planer/1182736860619>

Harnad, S. et al. (Sept. 2004). *The Access/Impact Problem and the Green and Gold Roads to Open Access*. Article in *Serials Review*, Vol. 30, No. 4, s. 310-314. Lokalisert 15.10.08 på Verdensveven via OAIster: <http://eprints.ecs.soton.ac.uk/15852/2/serev-revised.pdf>

Harnad, S. (2008). *What is Open Access?* Lokalisert 10.11.08 på Verdensveven: <http://www.eprints.org/openaccess/>

Harvard University (2008). *Harvard Law School Library. Open Access and Scholarly Publishing at the Law School*. Lokalisert 14.05.09 på Verdensveven: <http://www.law.harvard.edu/library/faculty/open-access.html>

Hernes, S. (april 2008). *Norske vitenskaplige artikler 2005-2006 og institusjonsarkiv*. Lokalisert 12.09.08 på Verdensveven: [http://www.uhr.no/forskning/open\\_access](http://www.uhr.no/forskning/open_access) & [http://www.uhr.no/documents/Publisering\\_og\\_forlagspolicy\\_Open\\_Access.pdf](http://www.uhr.no/documents/Publisering_og_forlagspolicy_Open_Access.pdf)

Howard Hughes Medical Institute (2003). *Bethesda Statement on Open Access Publishing*. Lokalisert 13.10.08 på Verdensveven: <http://www.earlham.edu/~peters/fos/bethesda.htm>

Jakobsson A. (03.04.2003). *UiOs system for elektronisk publisering – strategisk plan for 2003-2007*. Funnet 28.04.09 på UiOs nettside om DUO - Digitale utgivelser ved UiO under overskriften Informasjon om DUO og lenke til Strategisk plan 2003-2007: [http://www.duo.uio.no/dokumentasjon/generelt/om\\_duo.html](http://www.duo.uio.no/dokumentasjon/generelt/om_duo.html)

Jones R., Andrew T., MacColl J. (2006). *The Institutional Repository*. Oxford: Chandos Publishing.

Lawrence, S. (2001). *Online or invisible? Free online availability substantially increases a paper's impact*. Article in *Nature*, Vol. 411, No. 6837, p.521. Lokalisert 15.10.08 på Verdensveven: <http://web.archive.org/web/20021011092543/www.neci.nec.com/~lawrence/papers/online-nature01/>

Lowry, R. (1999-2009 a). *Non-Parametric Alternative: The Friedman Test*, kapittel 15a. I: *Concepts and Applications of Inferential Statistics*. Lokalisert 05.06.09 på Verdensveven: <http://faculty.vassar.edu/lowry/webtext.html>

- Lowry, R. (1999-2009 b). *Procedures Applicable to Ordinal Data: The Friedman Test*. I: *VassarStats: Web Site for Statistics Computation (menyvalg Ordinal Data)*. Lokalisert 05.06.09 på Verdensveven: <http://faculty.vassar.edu/lowry/VassarStats.html>
- Lund, A.C. (10.03.09). *Forskning når ut til flere*. I: Aftenposten (Morgenutg. Kulturdelen). Oslo, s. 8.
- Norges Teknisk-vitenskaplige universitet (NTNU) (05.12.2008). *Elektronisk publisering ved NTNU*. Lokalisert 28.04.09 på Verdensveven: <http://www.ntnu.no/ub/diva/index.php>
- OpenDOAR (2008). *Directory of Open Access Repositories*. Lokalisert 01.09.08 på Verdensveven: <http://www.opendoar.org/>
- Open Society Institute (OSI) (2002). *Budapest Open Access Initiative*. Lokalisert 20.09.08 på Verdensveven: <http://www.soros.org/openaccess/read.shtml>
- Open Society Institute (OSI) (2008). *Open Society Institute and the Soros foundations network*. Lokalisert 20.09.08 på Verdensveven: <http://www.soros.org/about/overview>
- Pharo, Nils (01.03.08). *Institusjonelt arkiv ved Høgskolen i Oslo?* Arbeidsgruppe for utredning av elektronisk publisering og arkiv ved Høgskolen i Oslo. Oslo: HiO. Lokalisert 01.09.08 på Verdensveven ved å følge lenken i artikkelen "Fritt for alle"  
<http://www.hio.no/content/download/86519/687247/version/1/file/utredning+om+instiusjonelt+arkiv+ved+HiO.pdf>
- Ringdal, K. (2007). *Enhet og mangfold : Samfunnsvitenskaplig forskning og kvantitativ metode*. Bergen: Fagbokforlaget, 2. utg.
- Salo, Dorothea (11.12.2007). *Innkeeper at the Roach Motel*. University of Wisconsin, Madison, USA. Lokalisert 01.10.08 på Verdensveven via OAster: <http://digital.library.wisc.edu/1793/22088>
- SHERPA/RoMEO (2008). *Publisher copyright policies & self-archiving*. Lokalisert 01.09.08 på Verdensveven: <http://www.sherpa.ac.uk/romeo/>
- Sølvberg, Ingeborg Torvik (30.05.07). *Institusjonelt arkiv ved NTNU*. Rapport fra arbeidsgruppe. Skal NTNU ha et institusjonelt arkiv? Trondheim: NTNU. Lokalisert 01.09.08 på Verdensveven: [http://www.ub.ntnu.no/ubit/nyhet\\_vedlegg/Institusjonelt\\_Arkiv\\_ved\\_NTNU\\_2007.pdf](http://www.ub.ntnu.no/ubit/nyhet_vedlegg/Institusjonelt_Arkiv_ved_NTNU_2007.pdf)
- SwePub (2008). *En samlad ingång till all svensk vetenskaplig publicering*. Lokalisert 26.03.09 på Verdensveven: <http://www.ub.gu.se/swepub.se/index.html>
- Thomson Reuters (2009). *ISI Web of Knowledge*. Lokalisert 25.05.09 på Verdensveven: <http://isiwebofknowledge.com>
- Universitetet i Oslo (UiO) (2005). *Strategisk plan for Universitetet i Oslo 2005-2009*. Lokalisert 19.10.08 på Verdensveven: [http://www.uio.no/om\\_uio/strategiskplan/2005-2009/](http://www.uio.no/om_uio/strategiskplan/2005-2009/)

Universitetet i Oslo (UiO) (13.10.2008). *Informasjon om DUO : Open Access-dag 14.oktober*. Lokalisert 14.10.08 på Verdensveven: <http://wo.uio.no/as/WebObjects/nettlogg.woa/wa/melding?id=95658&logg=1961>

Universitetet i Oslo (UiO) (2009). *NORA – Norwegian Open Research Archives. Avansert søk*. Lokalisert 01.10.08 på Verdensveven: <http://www.ub.uio.no/nora/oaister/topic.html>

Universitetet i Tromsø (UiT) (10.09.2008a). *NORA – Norwegian Open Research Archives*. Lokalisert 01.10.08 på Verdensveven: <http://www.ub.uit.no/wiki/openaccess/index.php/NORA>

Universitetet i Tromsø (UiT) (10.09.2008b). *Openaccess.no – om fri tilgang til vitenskaplige resultater*. Lokalisert 19.10.08 på Verdensveven: <http://www.ub.uit.no/wiki/openaccess/index.php/Openaccess.no:Om>

Universitets- og høyskolerådet (UHR) (25.01.2005). *UHR-anmodning*. [http://www.uhr.no/documents/05\\_29\\_Brev\\_medl.inst.pne\\_arkiver.doc](http://www.uhr.no/documents/05_29_Brev_medl.inst.pne_arkiver.doc). Funnet 12.09.08 på UHRs nettside [http://www.uhr.no/forskning/open\\_access](http://www.uhr.no/forskning/open_access) under overskriften ”UHRs aktiviteter”.

Universitets- og høyskolerådet (UHR) (30.01.2009). *Økt tilgang til norske forskningsresultater. Anbefalinger fra arbeidsgruppe oppnevnt av UHRs forskningsutvalg*. Lokalisert 28.04.09 på Verdensveven: [http://www.uhr.no/documents/Rapport\\_fra\\_Open\\_Access\\_arbeidsgruppe.pdf](http://www.uhr.no/documents/Rapport_fra_Open_Access_arbeidsgruppe.pdf)

Uppsala universitet (2009). *DiVA portal – Digitale Vitenskaplige Arkivet*. Lokalisert 02.05.09 på Verdensveven: <http://www.diva-portal.org/smash/search.jsf>

## 8. Vedlegg

### 8.1 Ordnet liste med oversikt over forfattere og deres artikkelproduksjon i NORA/DiVA

Listen gir oversikt over fordeling av antall artikler pr forfatter for emnekategori MAT fra UiB. Dersom en artikkel har flere forfattere med hovedforfatter/medforfatter vil artikkelen fremkomme i opptellingen for hver av forfatterne av artikkelen.

I eksempelet er forfatterne merket med stjerne (\*) knyttet til emnekategori MAT for UiB de som ble grunnlaget for videre søk i ISI.

Forfatter	Dok	Fak	FAKTA				
Birks Harry John Betteley	26	MAT *	Kilde				
Birks Hilary H.	9	MAT *	Dato				
Flower Roger J.	7	MAT *	IA				
Ramdani Mohammed	7	MAT *	Emnekategori				
Appleby P. G.	6	MAT *	Tidsperiode				
Fathi Adel A.	6	MAT *	Antall dok			62	
Peglar Sylvia M.	6	MAT *	Antall forfattere			179	
Rose Neil L.	6	MAT *					
Kraïem Mohammed M.	6	MAT *	Snitt ant forf pr dok			2,9	
Totland Ørjan	5	MAT *	Snitt ant dok pr forf			0,3	
Bjune Anne Elisabeth	4	MAT *					
Heegaard Einar	5	MAT *	<b>Fordeling forf/medforf og ant dok</b>	Ant forf	Ant dok	%	Akkum
Jones Vivienne J.	4	MAT *		1	26	0,6	
Patrick Simon T.	4	MAT *	Tabellen til høyre leses:	1	9	0,6	2
Elkhiati Najat	3	MAT *	Det finnes 1 forfatter (Birks Harry John Betteley)	2	7	1,1	4
Telford Richard J.	3	MAT *	som er registrert som forfatter for	5	6	2,8	9
Vandvik Vigdis	4	MAT *	26 artikler i BORA.	2	5	1,1	11
Bakke Jostein	2	MAT *	Det finnes 2 forfattere (Flower & Ramdani)	4	4	2,2	15
Barnekow Lena	2	MAT *	som er registrert som forfatter for	2	3	1,1	17
Bergmann Ørjan	2	MAT *	7 artikler.	19	2	10,6	36
Blom Tom	2	MAT *		143	1	79,9	179
Boyle John B.	2	MAT *		179		100	
Chan Tony F.	2	MAT *	* = Valgt for videre søk i ISI				
Christiansen Oddvar	2	MAT *					
.... forts.							

## 8.2 Forfatterutvalget - utgangspunkt for søk i ISI

Listen gir oversikt over hvordan søkene mot ISI ble satt opp og resultatene i form av antall artikler som ble hentet. I eksempelet viser jeg oppsett av søk og resultater for forfatterne knyttet til emnekategorien MAT for UiB.

Ant	Citation Report - oppsett av søk i ISI pr forfatter	Articles hentet	h-index	Articles plukket	Match
1	Author=(BIRKS HJB OR BIRKS HJBB) AND Address=(Norway) AND Document Type=(Article)	73	25	66	20
2	Author=(BIRKS HH) AND Address=(Norway) AND Document Type=(Article)	21	13	19	0
3	Author=(FLOWER R OR FLOWER RJ) AND Address=(Norway) AND Document Type=(Article)	2	2	2	0
4	Author=(RAMDANI M) AND Document Type=(Article)	26	3	21	0
5	Author=(APPLEBY P OR APPLEBY PG) AND Address=(Norway) AND Document Type=(Article)	14	11	14	1
6	Author=(FATHI A OR FATHI AA) AND Address=(Norway) AND Document Type=(Article)	1	1	1	0
7	Author=(PEGLAR S OR PEGLAR SM) AND Address=(Norway) AND Document Type=(Article)	8	7	8	2
8	Author=(ROSE N OR ROSE NL) AND Address=(Norway) AND Document Type=(Article)	9	8	9	4
9	Author=(KRAIEM M OR KRAIEM MM) AND Document Type=(Article)	8	4	3	0
10	Author=(TOTLAND O) AND Address=(Norway) AND Document Type=(Article)	40	13	34	5
11	Author=(BJUNE A OR BJUNE AE) AND Address=(Norway) AND Document Type=(Article)	12	5	9	2
12	Author=(HEEGAARD E*) AND Address=(Norway) AND Document Type=(Article)	27	10	24	5
13	Author=(JONES V OR JONES VJ) AND Address=(Norway) AND Document Type=(Article)	11	9	10	4
14	Author=(PATRICK S OR PATRICK ST) AND Address=(Norway) AND Document Type=(Article)	1	1	1	0
15	Author=(ELKHIATI N) AND Document Type=(Article)	4	2	2	0
16	Author=(TELFORD R OR TELFORD RJ) AND Address=(Norway) AND Document Type=(Article)	15	9	15	3
17	Author=(VANDVIK V) AND Address=(Norway) AND Document Type=(Article)	17	7	16	4
18	Author=(BAKKE J*) AND Address=(Norway) AND Document Type=(Article)	53	12	52	2
19	Author=(BARNEKOW L) AND Address=(Norway) AND Document Type=(Article)	2	2	2	2
20	Author=(BERGMANN O*) AND Document Type=(Article)	15	7	15	0
21	Author=(BLOM T*) AND Address=(Norway) AND Document Type=(Article)	1	1	1	1
22	Author=(BOYLE J) AND Document Type=(Article)	115	21	112	0
23	Author=(CHAN T OR CHAN TF) AND Address=(Norway) AND Document Type=(Article)	4	4	4	0
24	Author=(CHRISTIANSEN O*) AND Address=(Norway) AND Document Type=(Article)	15	9	14	0
25	Author=(E OR EDWARDS T OR EDWARDS TW OR EDWARDS TWD) AND Address=(Norway) AND Document Type=(Article)	2	2	2	2
26	Author=(HAMMARLUND D) AND Address=(Norway) AND Document Type=(Article)	4	3	3	2
27	Author=(KORHOLA A*) AND Address=(Norway) AND Document Type=(Article)	6	5	6	1
28	Author=(LEE TM) AND Document Type=(Article)	212	27	203	0
29	Author=(LIE J) AND Address=(Norway) AND Document Type=(Article)	5	3	5	0
30	Author=(MEZE HAUSKEN E) AND Address=(Norway) AND Document Type=(Article)	4	1	4	2
31	Author=(MYKLESTAD A) AND Address=(Norway) AND Document Type=(Article)	4	3	4	1
32	Author=(NESJE A) AND Address=(Norway) AND Document Type=(Article)	44	16	43	2
33	Author=(OLANDER H*) AND Document Type=(Article)	7	5	7	1
34	Author=(SINHA U) AND Document Type=(Article)	56	14	55	0
35	Author=(SOLHOY T) AND Address=(Norway) AND Document Type=(Article)	10	6	10	2
36	Author=(SAETERSDAL M) AND Address=(Norway) AND Document Type=(Article)	12	9	12	1
		<b>Opprinnelig tatt ut fra ISI</b>		<b>** Fjernet 08-09</b>	
		<b>Articles</b>	<b>h-index</b>	<b>Articles</b>	<b>Match</b>
Sum ant. artikler		860	280	808	69
Sum ant. forfattere		36		36	
Snitt		23,9	7,8	22,4	1,9
Citation Report - eksempel					
Author=(JONES V OR JONES VJ) AND Address=(Norway) AND Document Type=(Article)					
Timespan=1997-2009. Databases=SCI-EXPANDED, SSCI, A&HCI.					
*) RØD = "Article" ifølge BORA, men ikke "Peer reviewed; Journal article".					
**) Totalt 11 dok i ISI => 11 - 1 fra 2009 = 10 dok. NB! ISI-analysene baserer seg på totaltallene.					
Satte opp tidsrommet 1997 - 2009.					
Satte først tilår 2007. Enkeltiteringene for 08-09 ble borte, mens de fortsatt var akkumulert i "Total Citation"					

### 8.3 Artikkelutvalget - uttrekk fra ISI

Her viser jeg et eksempel på uttrekket av artikkel- og siteringsdata fra ISI slik det ser ut for én forfatter. Tabellen gir kun et utsnitt av dataelementene som blir hentet, men viser de dataene som brukes videre for beregningene i analysene. Alle artiklene for alle forfatterne utgjør artikkelutvalget i undersøkelsen. Den nederste tabellen viser sammendrag av data for forfatteren som akkumuleres til nivåene over for emnekategori, institusjon og land.

I eksempelet viser jeg uttrekket for en forfatter knyttet til emne kategorien MAT for UiB.

Title	Pub Date	Pub Year	Total Cit	IA	Average per Year	Year										SUM				
						1997	1998	1999	2000	2001	2002	2003	2004	2005	2006		2007	2008		
Environmental	JAN	2006	7		2,3											1	4	2	7	
Holocene	MAR	2006	5		1,7											1	2	2	5	
Spatial	FEB	2006	8		2,7											1	4	3	8	
A multi-proxy	JUN	2005	10		2,5										0	1	6	3	10	30
Recent	MAY	2004	12	1	2,4								8		0	1	0	3	12	
Quantitative	.	2002	63	1	9,0						0	5	8		16	14	8	12	63	
Plant species	.	1999	29		2,9			0	1	1	5	4	3		1	3	4	7	29	
Predicting	JAN	1998	20		1,8			1	1	1	2	2	1		2	2	2	4	20	124
			154	2		0	1	1	2	3	7	11	20		19	24	30	36	154	
													45					109		

TOT		Ant sit		Snitt	
Artikler i ISI **)	8				
Før IA	4			50,0	% av ISI-art publisert før IA, mens
Etter IA	4			50,0	% av ISI-art er publisert etter IA
Artikler i NORA/IA *)	6				
Før IA	6			100,0	% av NORA/IA-art publisert før IA, mens
Etter IA	0			0,0	% av NORA/IA-art er publisert etter IA
Artikler samkjørt	2				
Før IA	2			33,3	% av NORA/IA-art finnes også i ISI, men disse utgjør kun
Etter IA	0			25,0	% av forfatterens artikler funnet i ISI
ISI h-index	7				
Siteringer ALLE	154			19	%
Alle: 97-04	45	Alle: 97-04		6	Alle: 97-04 29,2
Alle: 05-08	109	Alle: 05-08		14	Alle: 05-08 70,8
Siteringer FØR	124			31	
FØR: 97-04	45	FØR: 97-04		11	FØR: 97-04 29,2
FØR: 05-08	79	FØR: 05-08		20	FØR: 05-08 51,3
ETTER: 05-08	30	ETTER: 05-08		8	ETTER: 05-08 19,5
Citation Report					
Author=(PEGLAR S OR PEGLAR SM) AND Address=(Norway) AND Document Type=(Article)					
Timespan=1997-2009. Databases=SCI-EXPANDED, SSCI, A&HCI.					
*) RØD = "Article" ifølge BORA, men er ikke "Peer reviewed; Journal article"!					
**) Totalt 8 dok i ISI og NB! ISI-analysene baserer seg på totaltallene.					
Satte opp tidsrommet 1997 - 2009.					
Satte først tilår 2007. Enkelt siteringene for 08-09 ble borte, mens de fortsatt var akkumulert i "Total Citation"					



## 8.4 Sammenstilling av uttrekk fra NORA/DiVA og ISI

Dette er aggregerte data som viser uttrekket fra NORA/DiVA og ISI, samt relasjonene mellom dataene. Noe er brukt direkte i analysene og ellers er dataene brukt videre i ulike beregninger for å finne økning/reduksjon i antall siteringer, andeler av og forskjeller mellom siteringer for IA og ikke-IA artikler.

<b>NORGE - MAT/MED/SAMF</b>					<b>SVERIGE - MAT/MED/SAMF</b>				
Antall forfattere	118				173				
	<b>Total</b>	<b>Total</b>	<b>Total</b>	<b>Total</b>	<b>Total</b>	<b>Total</b>	<b>Total</b>	<b>Total</b>	
	Tall	Pr forf			Tall	Pr forf			
<b>Artikler i ISI 1997-2007</b>	2111	18			8999	52			
Før IA	1128	10	53,4	% av ISI-art publisert før IA, mens	3497	20	38,9		
Etter IA	981	8	46,5	% av ISI-art er publisert etter IA	5442	31	60,5		
<b>Artikler i NORA 1997-2007</b>	321	3			2547	15			
Før IA	143	1	44,5	% av IA-art publisert før IA, mens	818	5	32,1		
Etter IA	178	2	55,5	% av IA-art er publisert etter IA	1726	10	67,8		
<b>Artikler med match i ISI og NORA</b>	156	1			1278	7			
Før IA	86	1	48,6	% av IA-art finnes også i ISI, men utgjør kun	502	3	50,2		
Etter IA	70	1	7,4	% av forfatterens artikler funnet i ISI	776	4	14,2		
			Snitt	%			Snitt	%	
<b>Siteringer ALLE</b>	33358	283	16		143727	831	16		
<b>Alle: 97-IA</b>	9489	80	4	28,4	21900	127	2	15,2	
<b>Alle: IA-08</b>	23869	202	11	71,6	121827	704	14	84,8	
<b>Siteringer FØR</b>	25416	215	23		88447	511	25		
<b>FØR: 97-IA</b>	9489	80	8	28,4	21900	127	6	15,2	
<b>FØR: IA-08</b>	15927	135	14	47,7	66735	386	19	46,4	
<b>ETTER: IA-08</b>	7942	67	8	23,8	55226	319	10	38,4	