

Humanistisk forskning og publisering i en digital kontekst

Europeiske filosofimiljøer trekker veksler på semantisk metadata-tagging

Alois Pichler & Gjermund Lanestedt, Unifob/Universitetet i Bergen & Teleplan AS

With the Wittgenstein Archives at the University of Bergen and the EU eContentplus funded DISCOVERY project as its touch stones, this article explores changed conditions and possibilities for digital editorial work, research and publishing in the humanities. In this article, emphasis is put on the dynamic and interactive aspects of digital editing, as well as the new opportunities which semantic web thinking and technology offer. Markup / text encoding is here considered key for both, as well as being crucial for adequately exploiting the possibilities offered by the digital turn in the humanities and higher education more generally. Additionally, the article promotes the strengthening and further development of open access and collaborative research models as beneficial for the humanities. The DISCOVERY project is described in some detail: Its importance lies in bringing the above dimensions together by aiming at giving open, multilingual access to a substantial digital resource of European philosophy sources, including multimedia, while at the same time being open for user specific filtering, semantic web processing and collaborative researching and learning.

Keywords: editorial theory, philology, ontologies, markup, open access, collaboration, digitisation, humanities computing, scholarly editions, TEI, versioning, XML

EU og digitale humaniora

I flere hundre år har penn og papir vært hovedredskap for humanistisk forskning.¹ Nå har det digitale mediet gjort sitt inntog også på dette feltet, og humaniora blir i stadig større grad "digitale" humaniora. Humaniora i dag benytter seg ikke bare av det digitale mediet, det digitale har selv blitt forskningsobjekt for humaniora.

EU peker på humanistisk forskning og kulturarven som sentrale objekter for digital teknologi og anser digital teknologi som et viktig redskap for en mer strategisk utnyttelse av Europas unike og rike kulturarv. Effektiv tilgang til og gjenbruk av digitale informasjonsressurser innen utdanning, forskning, innovasjon og produktutvikling oppfattes som en nøkkelfaktor i den voksende europeiske kunnskaps- og tjenesteøkonomien. EU ser Wittgensteinarkivet i Bergen (se <http://wab.aksis.uib.no>) som et eksempel på infrastruktur for samfunns- og humanistisk forskning, eller såkalt Research Infrastructure. Wittgensteinarkivet har i 2000 i samarbeid med Oxford University Press publisert en CD-utgave som inneholder hele Wittgensteins etterlatte originalmanuskripter, hans "*Nachlass*". Fra begynnelsen av har dette prosjektet vært orientert mot den digitale konteksten (se Huitfeldt 2001). Wittgensteinarkivets edisjonsfilologiske arbeid med Wittgensteins *Nachlass* har fått internasjonal oppmerksomhet og dets utgave blir av mange betraktet som eksemplarisk når det gjelder elektronisk vitenskapelig publisering. Wittgensteinarkivets arbeid har tidligere blitt omtalt i *Human IT* (se Cripps & Ore 1997). I dag tilhører Wittgensteinarkivet Avdeling for kultur, språk og informasjonsteknologi (AKSIS, <http://www.aksis.uib.no>) ved Universitetsforskning Bergen AS (Unifob AS), som er sterkt engasjert i digitale ressurser innen humaniora og har et fremtredende tekstteknologisk miljø. To tekstteknologiprojekter ved AKSIS, Medieval Nordic Text Archive (Menota, <http://www.menota.org>) og Wittgensteinarkivet, har nylig blitt positivt fremhevet og omtalt i en EU-rapport over europeisk forskningsinfrastruktur (se <http://www.aksis.uib.no/news/183>).

Mens Wittgensteinarkivet arbeidet med digital vitenskapelig utgivelse av Wittgensteins skrifter, har et tysk-fransk prosjekt samtidig arbeidet med å bygge et Web-basert forsknings- og læringsfelleskap rundt Friedrich Nietzsches tekster. Dette er HyperNietzsche-prosjektet som ble eta-

blert i 2001 (se <<http://www.hypernietzsche.org/>>), og som fra starten har forpliktet seg til å benytte åpen publisering i sitt arbeid. I løpet av seks år har HyperNietzsche, etter en lisensavtale med arkiv og innholdseiere, gjort fritt tilgjengelig mer enn 3 000 digitale faksimiler av sider fra Nietzsche-materiale og familiefotografier (se D'Iorio 2001). I sitt arbeid har HyperNietzsche-prosjektet hele veien samarbeidet tett med den italienske IT-bedriften Net7 (se <<http://www.netseven.it/>>) som har stått for utviklingen av teknologiplattformen til prosjektet.

HyperNietzsche, Net7 og Wittgensteinarkivet har gått sammen og – sammen med andre partnere – dannet et konsortium som står bak DISCOVERY-prosjektet (se <<http://www.discovery-project.eu/>>), som har fått støtte av EUs eContentplus-program. DISCOVERY går ut på å gjøre ca 50 000 sider filosofitekster og flere timer med lyd- og bildeopptak fra europeisk filosofihistorie fritt tilgjengelig for forskning og læring på Web. Som i HyperNietzsche vil man bygge opp et forsknings- og læringsfelleskap og et publiseringssystem rundt disse ressursene. Som i Wittgensteinarkivets arbeid vil man for primærtekstene bruke de mest adekvate redskapene og metodene fra digital utgivelsesfilologi. I tillegg vil man utvikle søkeredskaper som muliggjør ”semantisk søk”, og anvender og videreutvikler til dette formålet ”Semantic Web”-tenkning og -teknologi, og semantisk koding mer generelt.

I denne artikkelen skal vi først beskrive eContentplus-programmet litt nærmere og diskutere dets betydning for digitale humaniora. Så presenterer vi DISCOVERY-prosjektet i hovedtrekk. Til slutt presenterer vi Wittgensteinarkivets bidrag til prosjektet og bruker dette for å drøfte nøkkelaspekter ved humanioras arbeid i den digitale konteksten.

eContentplus-programmet

EU-kommisjonen er opptatt av IT som virkemiddel og driver i økonomien. Området digitalt innhold og tilgang (til innhold) står sentralt. Kultursektoren og humaniora står særlig sentralt i denne sammenheng. Innenfor det nylig lanserte syvende rammeprogram for forskning (FP7) er det betydelig fokus på hva IT kan bidra med nettopp på dette området. I sin første utlysning i 2006 lyste FP7 ut hele 102 millioner euro til forskning på området ”Digital Libraries and Content”, som en av 7 ”hovedutfordringer” innenfor rammeprogrammet. Ytterligere 100 mil-

lions euro lyses etter planen ut senere i år. Semantiske søkesystemer, flerspråklighet og mer effektive systemer for digital bevaring er helt sentrale temaer innen dette området.

Parallelt med den mer langsiktige forskningsinnsats på området er EU opptatt av at det raskt etableres løsninger som gjør det enklere å få tilgang til digitalt innhold på tvers av land, språk og kulturgrenser. Vår digitale kulturarv står sentralt i denne sammenheng, og EU-kommisjonen har et eget såkalt "flaggskip"-initiativ kalt "Digital Libraries Initiative" – eller "Europe's Cultural and Scientific Heritage at a Click of a Mouse" som det også omtales. Dette handler om mer enn mulighetene for fødererte søk på tvers av bibliotekdatabaser. Det handler om enklere tilgang til innhold i alle typer samlinger, museer, repositorier, arkiv, etc. Enklere tilgang til digitalt innhold vil gjøre det mulig for flere å ta del i den europeiske kulturarven, og vil gi viktige læringsressurser for skole og utdanning. Enklere tilgang til innhold vil også stimulere forskning og bidra til utviklingen av innovative, verdiøkte digitale tjenester.

EU etablerte allerede i 2001 et eget innholdsprogram (eContent) som var innrettet på å finansiere gode prosjekter på innholdsområdet. Etterfølgeren eContentplus (se <<http://ec.europa.eu/econtentplus>>) er et 4-årig program (2005-2008) for å stimulere til bedre tilgang, økt bruk og nytte av digitalt innhold innen områder hvor markedsdynamikken er svak: geografisk informasjon, digitale læringsressurser, kulturelt innhold og vitenskapelig informasjon (forskningspublisering og -data). Programmet har et budsjett på 149 millioner euro (mer enn 1,1 milliarder kroner) som støtte til prosjekter som håndterer organisatoriske barrierer og som ved hjelp av best tilgjengelig teknologi forbedrer tilgangen og anvendeligheten til digitalt materiale i en flerspråklig kontekst.

eContentplus-programmet materialiserer EUs visjoner på området. I årets utlysning sies det eksempelvis at de digitale samlingene nå beveger seg inn i online-omgivelser, og derved bidrar til europeisk tilstedeværelse i det 21. århundres opplevelsesøkonomi. Digitalt kulturelt og vitenskapelig innhold muliggjør framveksten av nye verdiøkte tjenester innen forskning, læring og underholdning. Og framfor alt gis folk flest tilgang til Europas rike kulturarv på en enkel måte. Som eksempel på det siste vil det eContentplus-finansierte prosjektet European Digital Library (EDL) gjøre tilgjengelig minst 6 millioner bøker, dokumenter og annet materi-

ale innen utgangen av 2010. Målet er ett enkelt tilgangspunkt for innholdsdata-baser i museer, samlinger og andre kulturinstitusjoner i et stort antall europeiske land.

eContentplus vil også bidra til større spredning av forskningsrelatert innhold, slik som artikler, posters, proceedings, monografier, kompendier og underliggende datasett. Gjennom målrettede forsøk med åpen tilgang til vitenskapelige innholdsdata-baser, ønsker EU å utfordre de tradisjonelle fagtidsskriftenes hegemoni når det gjelder forskningspublisering. De ser for seg en rekke prosjekter hvor det eksperimenteres med nye forretningsmodeller og publiseringsprosesser. Berørte aktører er akademia, biblioteker, U&H-institusjoner, forlag og de som finansierer forskningen. EUs mål er at det skal bli vesentlig enklere å gjenfinne, sette sammen og aggregere forskningsrelatert innhold på tvers av land.

Blant mange søknader har så langt 36 prosjekter sluppet igjennom nåløyet til eContentplus-programmet. Norge deltar i 10 av disse, herunder i DISCOVERY-prosjektet. DISCOVERY vil skape mer tilgang og merverdi til filosofiske tekster og annet digitalt, forskningsbasert innhold om filosofi. Prosjektet samsvarer derfor svært godt med EUs "cyberfilosofi". DISCOVERY er i størrelse og varighet ganske typisk for eContentplus-prosjektene, med et totalbudsjett på ca NOK 20 millioner, og med 3 års varighet.

DISCOVERY-prosjektet

I 2004 inngikk HyperNietzsche-prosjektet et samarbeid med Wittgensteinarkivet og flere andre europeiske arkiv og utgivelsesprosjekter om en "Groupement de Recherche Européen plus *Hyper-Learning*" (GDRE+, <<http://www.hyperl.org/index.php?page=/hyper/home>>). Denne Hyper-Learning-gruppen ledes av franske Centre National de la Recherche Scientifique (CNRS) og har som erklært formål å bygge et digitalt forsknings- og læringsfelleskap rundt et utvalg europeiske forfattere. Prosjektet tangerer en rekke områder og kompetanser, bl.a. utgivelsesfilologi, elektronisk publisering, tekstkoding, e-læring og håndtering av IPR. I 2005 fikk et konsortium dannet av medlemmer av gruppen finansiell støtte fra det europeiske COST office til COST action A32 ("Open Scholarly Communities on the Web"). COST action A32 pågår fra 2006 til 2010 og har to hovedformål:

1. “to create a digital infrastructure for collaborative humanities research on the Web”;
2. “to establish and foster the growth of Scholarly Communities that will provide feedback to the IT developers regarding the needs and expectations of humanities researchers and will serve as a core group of early adopters.” (<<http://www.cost-a32.eu/index.html>>)

Siden starten har konsortiet i COST action A32 vokst til å inkludere en stor mengde prosjekter og partnere fra 13 europeiske land.

DISCOVERY-prosjektet er i stor grad et resultat av det mangeårige samarbeidet i denne Hyper-Learning gruppen og har igjen Hyper-Nietzsche, Net7 og Wittgensteinarkivet som sentrale aktører. I korte trekk går DISCOVERY ut på å etablere en omfattende, strukturelt og semantisk tagget innholdsressurs for forskning og undervisning i filosofi. Flere digitale filosofiske arkiver og -portaler knyttes sammen, og det skal på plass applikasjoner som lar brukeren stille avanserte spørsmål inn mot tekstene – om filosofiske emner på tvers av alle tilgjengelige kilder. Dette gir muligheten til å teste resonnementer og derigjennom etablere helt ny kunnskap. Forskernes funn beriker igjen arkivene med nytt innhold og nye, kontekstuelle metadata. De metadata som benyttes, er definert i felles ontologier eller ”begrepssammenhenger”, noe som igjen bidrar til å øke nytteverdien til innholdsressursene. Applikasjonene vil være tilgjengelig på minst 6 nasjonalspråk, deriblant norsk. DISCOVERY etablerer altså et rammeverk for flerspråklig håndtering og semantisk beskrivelse av digitalt filosofisk material. Teknisk skal løsningen baseres på open-source teknologien *Hyper* som tillater dynamisk, kontekstuell lenking av dokumenter, og peer-to-peer løsningen *dbin* (se Barbera & Giomi 2003, Barbera 2006 og <<http://www.dbin.org/>>).

Franske CNRS vil bidra med 39 000 sider digitalt Nietzsche-materiale. Italienske Consiglio Nazionale delle Ricerche (CNR) deltar med 12 000 tekstsider fra de gamle greske filosofer og filosofiske tekster fra det 16., 17. og 18. århundre. Wittgensteinarkivet bidrar med 5 000 sider Wittgenstein-materiale. Innholdet er mye tekst, men ikke bare. I tillegg til alt dette tekstlige materialet stiller italienske RAI Net med sine

store arkiv av intervjuer, reportasjer og ulike typer multimediainnhold knyttet til filosofiske spørsmål.

Ett av DISCOVERY-prosjektets mål er å vise at kommersiell og åpen publisering lever godt sammen og faktisk trenger hverandre. Dermed vil prosjektet bidra til at den vitenskapelige publiseringsarenaen utvikles i en antatt riktig retning. Gjennom DISCOVERY vil forskersamfunnet kunne berike filosofiressurser og øke deres verdi ytterligere. De kvalitets-sikrede delene av ressursen vil bli gjort fritt tilgjengelig for alle og skal vedlikeholdes også etter prosjektets slutt. Forlag kan foreta ytterligere investeringer i ressursene og bearbeide dem videre til utgaver. Deler av inntektene fra disse utgavene vil kunne flyte tilbake til ressurseierne og kan så igjen bidra til å vedlikeholde og videreutvikle dem. Det er gode grunner til å tro at åpen og kommersiell publisering slik kan komplettere hverandre.

Wittgensteinarkivet i DISCOVERY

Wittgensteinarkivet har vært en viktig aktør i GDRE+ Hyper-Learning (2004), i COST action A32 (2004) og også i DISCOVERY-prosjektet (2005). Det har først og fremst vært dets status som europeisk forskningsinfrastruktur i humaniora, dets erfaring med tekstkoding og vitenskapelig digital utgivelse, dets forvaltning av Wittgensteins *Nachlass* i digital form og dets sterke engasjement i digitale humaniora mer generelt som har gjort Wittgensteinarkivet til en attraktiv partner og viktig bidragsyter i disse sammenhenger. I det følgende presenterer vi Wittgensteinarkivets bidrag til DISCOVERY i større detalj, noe som gir muligheten til å diskutere noen hovedutfordringer for digitale humaniora mer inngående.

Vitenskapelig utgivelse

Hva skjer med vitenskapelig utgivelse i den digitale kontekst? Det er mye som taler for at det tradisjonelle konseptet hvor en utgivelse er endelig og statisk, er på vei til å bli historie. Dette skyldes digitalitetens inntog også i utgivelsesfilologien. I Norden er det et livlig og bevisst utgivelsesmiljø som har diskutert digital utgivelse over lang tid i bl.a. Nordisk nettverk for edisjonsfilologi (se <<http://www.nnedit.org>>). Ved Henrik Ibsens Skrifter (HIS, <<http://www.ibsen.uio.no/his/>>), Søren Kierkegaard

Forskningscenteret (SKF, <<http://www.sk.ku.dk>>), Medieval Nordic Text Archive (Menota, <<http://www.menota.org>>) og Wittgensteinarkivet har man noen av de mest avanserte vitenskapelige utgivelsesprosjekter, som bevisst forholder seg til de nye mulighetene digitaliseringen gir, men også utfordringene som kommer med den. Man har i disse prosjektene kommet til at tekstkoding eller *markup* er helt uunnværlig for forberedelse og gjennomføring av vitenskapelige utgaver.

Ved AKSIS i Bergen, som bl.a. omfatter Wittgensteinarkivet og Menota, er man opptatt av å føye tradisjonell edisjonsfilologisk metode og praksis inn i den digitale kontekst. I denne sammenheng har man utviklet begrepene "dynamisk utgivelse" og "editio continua" (se Pichler & Haugen 2005). Begrepene uttrykker at vitenskapelig utgivelse nå oppfattes som dynamisk og "open-ended". Utgivelse er en uavsluttet prosess, hvor utgaver og utgivelsespremisser løpende må vedlikeholdes, revideres og forbedres i forhold til ny innsikt i tekstsituasjonen, tekniske krav og brukerbehov. Der vi før hadde en mer eller mindre statisk og avsluttet tekst, vil det nå eksistere mange tekster, som alle kan være like autentiske som den som tidligere var den ene rette. Og dersom utgaven baserer seg på en tekstkodet base, det som i Wittgensteinarkivets tilfelle kalles "maskinleselig versjon", vil brukeren i stadig større grad selv kunne bestemme utgavens omfang og egenart. En slik tenkning har også vært til stede i tradisjonell edisjonsfilologi, men vi tror det er først nå i den digitale konteksten at det virkelig materialiseres.

Går vi ett skritt lenger blir utgivelsen *interaktiv og dynamisk*. Dette kan for eksempel innebære at brukeren møter en webside som gir tilgang til et elektronisk tekstarkiv som består av kodete transkripsjoner, og til programvare som tillater å konvertere disse transkripsjonene til "web-utskrifter" som svarer til brukerens *hic et nunc* behov. Konkret kan det foregå via et webskjema hvor brukeren krysser av mellom flere valgmuligheter, for eksempel når det gjelder filtrering av strøket tekst. Skal den bli tatt med; hvordan skal den bli vist; eller skal kanskje bare en bestemt type overstrøket tekst komme med? Om slike valg blir mulige er selvfølgelig avhengig av om transkripsjonene er kodet slik at det tillater denne type filtrering. Wittgensteinarkivet har i 2000 publisert *Bergen Electronic Edition* (BEE) (se Huitfeldt 2001). BEE utnytter ikke fullt det potensiale og de differenseringsmuligheter som ligger i Wittgenstein-

arkivets maskinleselige versjon, og den tradisjonelle utgaven vil generelt ikke ta hensyn til behovet for slike detaljerte valg. Med interaktiv dynamisk utgivelse kan derimot et slikt behov dekkes, slik at brukeren kan overta utgivelsesfunksjoner og bestemme hva som skal komme med i utgaven og hvordan. Ved interaktiv dynamisk utgivelse får altså forholdet mellom teksten og brukeren så vel som forholdet mellom utgiveren og brukeren samhandlingskarakter, der det før var enveiskommunikasjon.

Kritiske røster hevder at en slik "dynamisk" og brukerorientert oppfatning av utgivelsens og utgavens rolle fører til ukontrollert relativisme. Det vil kanskje "konstrueres" vidt forskjellige tekster av det samme materiale, slik at man kanskje ikke lenger kan kommunisere med hverandre om teksten. Men dette blir motsagt av mange som har sitt profesjonelle virke på området. De hevder at det nettopp er tekstkodingen — som er et nøkkelement i digital dynamisk utgivelse og som tillater å dokumentere de enkelte utgave- og verkkonstruksjonene — som gir mulighet til å kommunisere effektivt med andre om innholdet. Det har nå i lengre tid gått en diskusjon om realisme/anti-realisme/relativisme i tekstkoding og utgivelse (se for eksempel Biggs & Huitfeldt 1997, Eggert 2005). I denne diskusjonen har det realistiske og statiske tekstsynet som ligger til grunn for mye av tradisjonell edisjonsfilologi, blitt problematisert. Samtidig har diskusjonen vist at en anti-realistisk og dynamisk forståelse langt ifra fører til en vilkårlig og relativistisk utgivelsespraksis, men tvert imot gjør oss i stand til å beskrive vår utgivelses- og tekstpraksis adekvat – og ofte er bedre egnet som teoretisk grunnlag for utgivelse enn det realistiske synet (se Pichler 2002 og Pichler & Bruvik under utg.).

Wittgensteinarkivets BEE anses fortsatt som eksemplarisk på mange måter, både i mer filosofisk sammenheng og i generell edisjonsfilologisk sammenheng. Noen av sidene ved utgaven som kan virke problematiske er drøftet i Pichler & Haugen 2005, bl.a. hva gjelder diplomatarisk og normalisert utgivelse mer generelt. Et prinsipp BEE følger og som fortsatt kan anses som aktuelt, er formidling av materialet i tre versjoner som er lenket til hverandre på sidenivå: en diplomatarisk, en normalisert og en faksimileversjon. Man kan altså si at BEE består av tre utgaver som er sammenføyet til én "kombinert" utgave. Dette er bare mulig takket være tekstkoding som grunnlag for hele utgivelsesprosjektet. I perioden

1990-2000, da Wittgensteinarkivets maskinleselige versjon av *Nachlass* og BEE ble til, ble det brukt MECS WIT som kodespråk. MECS ble utviklet av Claus Huitfeldt på slutten av 1980-tallet, mens WIT-delen kom til på Wittgensteinarkivet under arbeidet med transkribering av Wittgensteins *Nachlass* (se Huitfeldt 2001). I dag er det XML TEI (P5) som er grunnlaget for utgivelse og bruk av materiale.

I Pichler & Haugen 2005 blir "dynamisk utgivelse" beskrevet som en videreutvikling av det kombinerte utgivelseskonseptet benyttet i BEE. Mens den kombinerte utgaven fortsatt fastholder på å levere et antall fikserte utgaver, byr interaktiv dynamisk utgivelse også på muligheten for å la brukeren selv "bygge" sin egen utgave for "her og nå"-behovet. Dette er mulig ved at utgiveren stiller til disposisjon for brukeren det kodete materialet og programvare som gjør det mulig for brukeren å bearbeide materialet slik at det samsvarer med hans spesifikke behov. Men en slik utgivelsesmodell ekskluderer ikke at det samtidig også kan tilbys fikserte utgaver à la BEE. Disse har en funksjon som faste holdepunkter og er tilrettelagt for de mest representative brukssituasjonene.

Originalversjonen av BEE kommer på CD-ROM, utgitt av Oxford University Press (OUP). Det er fortsatt bare denne originalutgaven som inneholder kombinasjonen av den normaliserte, den diplomatariske og faksimileutgaven. Den normaliserte versjonen alene er imidlertid også gjort tilgjengelig på internett, utgitt av InteLex Corporation (<<http://www.nlx.com>>). I tillegg til dette har Wittgensteinarkivet gjort fritt tilgjengelig deler av *Nachlass* på sine websider, alltid med et øye på å teste videreutvikling av utgaven og alternative formater og opplegg. Dette innbefatter en normalisert og en diplomatarisk versjon av Wittgenstein *Nachlass*-manuskript 115 (ca 300 sider), som ble gjort tilgjengelig på Wittgensteinarkivets webside i 1996 (se <http://wab.aksis.uib.no/wab_115www.page>). Denne webutgaven kommer i to HTML-filer, som inneholder lenker til hverandre på sidenivå. I tillegg kan man bla i begge filene samtidig. Utformingen av den diplomatariske og den normaliserte versjonen er forskjellig fra utformingen av disse versjonene i BEE, men filene er selvfølgelig produsert fra samme datagrunnlag som BEE, nemlig den maskinleselige versjonen. Forskjellen kommer av at man bruker et annet filter hvor man har valgt å "realisere" visse koder på annet vis enn for BEE. Dette viser at disse to hovedformatene ikke er

absolutt fastlagte størrelser. Deres definisjon kan variere selv om man generelt kan si at den normaliserte versjonen først og fremst er tenkt for *tekst*-presentasjon og altså er rettet mer mot det konseptuelle, mens den diplomatariske har hovedfokus på *dokument*-presentasjon og altså er rettet mer mot det fysiske. I tillegg kan man på det samme nettstedet hvor man finner HTML-utskriftene av manuskript 115 også inspisere kildetranskripsjonene som disse HTML-utskriftene ble produsert fra.

Etter samme mønster er diplomatariske og normaliserte versjoner av typoskript 201a senere blitt publisert. Webpublikasjonene av *Nachlass*-delene 115 og 201a er fortsatt basert på MECS WIT-koden av den maskinleselige versjonen. I mellomtiden har Wittgensteinarkivet arbeidet med migrasjon av den maskinleselige versjonen til XML TEI-kode. Et tidlig eksempel på dens utforming og anvendelsesmuligheter er "Wittgenstein MS101 from September 1914", som først ble gjort tilgjengelig i 2001 (se <http://wab.aksis.uib.no/wab_sept1914.page>). "Wittgenstein MS101 from September 1914" publiserer en del av Wittgenstein *Nachlass*-manuskript 101 på Web både som HTML- og som PDF-utskrift. I tillegg er disse utskriftene lenket til faksimile og andre komplementære ressurser, som for eksempel korrespondanse. En viktig dimensjon ved dette prosjektet var å teste mulighetene for visning av spesialtegn på Web og i forskjellige nettlesere. Man kom frem til at det beste vil være å følge Unicode standard (<<http://unicode.org/>>), noe Wittgensteinarkivet fortsatt holder fast ved. En annen viktig innovasjon i dette prosjektet var at man for første gang kunne tilby brukerstyrt interaktiv dynamisk utgivelse. Funksjonen er fortsatt aktiv (<<http://wab.aksis.uib.no/sept1914/oxygen16/index.html>>), selv om den baserer seg på en annen XML TEI-kode enn den Wittgensteinarkivet har lagt seg på i dag (se <<http://wab.aksis.uib.no/transform/wab.php>>).

Samtidig med "Wittgenstein MS101 from September 1914" var Wittgensteinarkivet med i et annet, komplementært prosjekt, "Tracing Wittgenstein" (<<http://wittgenstein.philo.at/>>). Her har man føyet sammen diplomatariske og normaliserte versjoner av en del av *Nachlass*-manuskript 115 på den ene siden og en filosofisk kommentar til teksten på den andre. Primærteksten så vel som kommentaren er relatert til hverandre og er navigerbare på avsnittsnivå; dette var nytt i forhold til både de daværende webpublikasjonene (manuskript 101, manuskript

115, typoskript 201a) og BEE. Struktureringen av ressursen på avsnittsnivå har flere grunner: avsnittet er nemlig ikke bare en enhet som er veldig godt egnet som grunnenhet ut fra datapraktiske hensyn (for eksempel ved visning av søkerresultater), men faktisk også Wittgensteins primære refleksjons- og arbeidsenhet. I "Tracing Wittgenstein" ble det da også lagt vekt på å kunne forandre på rekkefølgen til de enkelte avsnittene, enten etter Wittgensteins egne instruksjoner for omrangering eller i samsvar med egne forskningsinteresser. Et resultat av prosjektet som integrerer de nevnte elementene kan lastes ned fra Wittgensteinarkivets webside (<http://wab.aksis.uib.no/wab_115ape.page>).

Måten å produsere disse webutskriftene på er standard ved AKSIS: transkripsjonen av teksten finnes i en kodet XML-fil, som blir sendt gjennom stilark som inneholder instruksjoner for konvertering av XML-kodene. Avhengig av om man vil produsere en diplomatarisk, normalisert eller annen utgave, velger man det stilarket/den delen av stilarket som inneholder instruksjonene for den respektive utgaven. Både XML-transkripsjonen og stilarket er reviderbare: dette gir stort utbytte når man vil forandre layout og visning for flere tusen sider bare ved å forandre én linje i stilarket. I DISCOVERY vil faksimile og "utgavene" være lenket til hverandre på avsnittsnivå; avsnittet er altså grunnenheten også i dette prosjektet. Dette arver DISCOVERY fra HyperNietzsche-prosjektet, men det føyer seg også godt til utviklingen ved Wittgensteinarkivet. Kommentarer, oversettelser, artikler og andre komplementære bidrag lenkes opp til det respektive avsnittet. Denne fremgangsmåten forutsetter at vi har et stabilt referansesystem ("sigler") for kildemateriale, noe som ble introdusert tidlig i prosjektet.

Wittgensteinarkivets bidrag til DISCOVERY ville vært utenkelig uten de ovennevnte miniprojektene, som testet og utviklet publisering av Wittgensteins *Nachlass* på Web, sammenlenking med og inkludering av komplementære ressurser, overføring av den maskinleselige versjonen til XML TEI, re-orientering av utgivelsesmodellen i retning av avsnitt istedenfor side, og brukerstyrt og interaktiv dynamisk utgivelse. For DISCOVERY skal Wittgensteinarkivet legge 5 000 sider Wittgenstein *Nachlass*-materiale ut på nettet, både i diplomatarisk, normalisert og faksimileutgave, men også i flere andre varianter, som inkluderer en "studieutgave" og et interaktivt grensesnitt hvor brukeren kan "krysse" av

for format og i grunnen bestemme selv hvordan de kodete kildetranskripsjonene skal konverteres.

Den interaktive dynamiske utgivelsesmodellen vil også inkludere en funksjon som er mer kjent fra tekstanalyse enn fra edisjonsfilologien. Den samme XML-teknologien som tillater en å filtrere ut tekst som er overstrøket eller fokusere på tekst som tilhører et bestemt skrivestadium kan nemlig også brukes til å fokusere på sitater, spørsmål, metaforisk tale og andre elementer som er interessante i analysen av innholdet. Dette vil vi se nærmere på i neste avsnitt, hvor vi tar for oss *semantisk koding*.

Semantisk koding

DISCOVERY-prosjektet illustrerer godt mange av de muligheter dagens humaniora og kulturvitenskap står ovenfor ved bruk av moderne informasjonsteknologi. Men mulighetene innebærer også nye utfordringer. En av disse er det åpenbare behovet for semantisk koding av alt kildemateriale for å muliggjøre effektive, semantiske søk i det omfattende materialet. Mens de fleste er vant til å benytte Google eller andre søkemotorer som baserer seg på indekseringer, gir semantisk koding av innholdet mulighet til å ta i bruk søke- og filtreringssystemer med atskillig større presisjon og anvendelighet til forskning og studier. Frie tekstsøk gjennom Wittgensteins *Nachlass* er allerede mulig i dag, gjennom BEE, men de gir i all hovedsak bare de ønskede resultater hvis de ordene man leter etter også finnes i teksten. Søk som er forberedt gjennom semantisk koding vil derimot la brukeren gjenfinne avsnitt som er relevante for et bestemt forskningsspørsmål, selv om de ordene som man forbinder spørsmålet med, ikke forekommer i selve teksten. Nytteverdien blir enda større hvis vi tenker oss dette anvendt på hele filosofiressursen som gjøres tilgjengelig i DISCOVERY. Når man så søker på et emne vil man få opp alle tekstene som omhandler emnet, og ikke bare de tekstene som faktisk bruker de ordene som beskriver emnet, på tvers av forskjellig termbruk hos forfatterne. På denne måten vil man kunne finne sammenhenger mellom tekster som ellers ville vært vanskelige å oppdage.

Semantisk koding ved Wittgensteinarkivet innebærer at man gir – i en separat XML-fil og i henhold til flere kategorier – hvert enkelt avsnitt i Wittgensteins *Nachlass* tematiske etiketter og beriker transkripsjonen

med kortfattet meta-informasjon om tekstens innhold. På denne måten får vi søkemuligheter vi ikke får med fritekstsøk. Det kan for eksempel være at Wittgenstein i et avsnitt snakker om "privatspråk"-problematikken uten å bruke ordet "privatspråk", men ved å diskutere for eksempel tannpine (se Hrachovec 2000). Hvis et slikt avsnitt kodes semantisk under "privatspråk" vil man kunne finne det hvis man leter etter alle drøftinger av "privatspråk", selv om selve ordet ikke brukes av Wittgenstein i denne sammenhengen.

Mulighetene å finne og samordne visse avsnitt blir enda større hvis alle semantiske etiketter er del av det som i datavitenskap kalles for en "ontologi". En ontologi er et hierarkisk begrepssystem som appliseres på et bestemt område, i vårt tilfelle Wittgensteins *Nachlass*-tekster. Den semantiske berikings tematiske etiketter relateres altså til hverandre på hierarkisk vis, og de semantiske sammenhengene formaliseres og sammenbindes på denne måten gjennom et "ontologisk" vev. Dermed blir det mulig å foreta avledninger, siden begrepene er hierarkisk relatert til hverandre og arver egenskaper fra deres overbegrep. Ved hjelp av programvare kan en avlede ny kunnskap og innsikt i tekstens struktur og innhold, og dets interne relasjoner. Samtidig er det nettopp den strenge og formelle oppbyggingen av en ontologi som gjør dens anvendelse på filosofiske tekster problematisk. Begrepsinnholdet til filosofiske tekster og humanioras objekter mer generelt lar seg ikke på samme måte hierarkisere som for eksempel begrepsområder i botanikk, juss eller medisin. Likevel finnes det mange fordeler ved "ontologisering" som man ugjerne vil gå glipp av (se Zöllner-Weber & Pichler 2007). Derfor innbefatter Wittgensteinarkivets bidrag til DISCOVERY også en semantisk koding av Wittgensteins *Nachlass* som arbeider ontologisk, og som også vil kunne relateres til en mer allmenn ontologi for filosofiske tekster. Den mindre formelle delen av den semantiske kodingen vil i mange tilfeller være mer nyttig for brukeren, mens den ontologiorienterte semantiske kodingen vil tillate søk og operasjoner på et veldig generelt nivå som også omfatter andre tekster enn Wittgensteins. Det er Department of Electronics, Artificial Intelligence and Telecommunications at the Polytechnic University of Marche, Ancona (DEIT, se <<http://www.deit.univpm.it/>>) som ivaretar den tekniske delen av ontologibyggingen for hele DISCOVERY-prosjektet.

I praksis ser Wittgensteinarkivets struktur for den semantiske kodingen (per august 2007) ut som følger. En mengde XML "list"-elementer i en separat XML-fil refererer til de respektive avsnittene i *Nachlass* som de inneholder semantiske etiketter for. Selve etiketteringen er inkludert i fem "item"-elementer. Det første elementet inneholder en gjengivelse av hovedpoenget/-poengene til avsnittet, gjerne i form av en kort setning eller en frase. Det andre elementet inneholder tematiske stikkord, om de selv forekommer i teksten eller ikke. Det tredje elementet *kan* inneholde en kort kommentar til avsnittet, som belyser avsnittet ytterligere, men samtidig har elementet den funksjonen at det gir evidens for etiketteringen i de første to elementene. Det fjerde elementet inneholder eller er en plassholder for den strenge ontologi-kodingen som skal brukes på dette avsnittet. Det femte elementet inneholder et filosofisk nøkkelord som passer på avsnittet, av den type man finner i filosofiske leksika. Brukerne vil dermed kunne undersøke behandlingen av bestemte filosofiske problemer hos en rekke forskjellige filosofer ved å følge tråden som er lagt inn ved etiketteringen med et filosofisk nøkkelord.

I tillegg kodes det i filen med kildetranskripsjonen visse stilistiske og form-elementer, for eksempel om noe er en dedikasjon, et forord, et spørsmål, et eksempel, direkte tale, sitat, en sammenligning, om det inneholder metaforisk bruk av ord, osv. Dette er elementer som en rekke Wittgenstein-forskere er opptatt av. Denne kodingen ligger altså i filen med kildetranskripsjonen, nøyaktig på samme måte som de øvrige transkripsjonskodene som identifiserer overstrøket tekst, innskudd, syntaktiske forhold, osv.

Den semantiske kodingen kan ekstraheres og samordnes med visning av tekst på forskjellige måter. "Hovedpoenget" og/eller stikkordene og nøkkelordene kan for eksempel filtreres ut og vises som "mouse over" ved avsnittet det refererer til; det kan gjengis i forskjellige språk slik at brukere som ikke behersker tekstens originalspråk så godt på denne måten får en første innføring i teksten. Men de kan også prosesseres uavhengig gjennom XQuery og lignende funksjoner.

Wittgensteins innholdsfortegnelse til hans "Big Typescript" (det "store maskinskrevne manuskriptet") fra 1932-1933 er en av de sjeldne innholdsfortegnelser han laget til sine skrifter. Det vanlige var at han skrev i korte avsnitt som var atskilt fra hverandre gjennom en eller flere

blanke linjer. Hans bemerkninger har derfor ofte blitt sammenlignet med aforismer. Vanligvis følger de ikke etter hverandre på en lineær og systematisk måte (som man ville forvente av et tradisjonelt akademisk filosofisk verk), men med brå innholdsmessige overganger mellom avsnittene. Mange lesere føler de mister retningen i hans tekster. "Big Typescript" regnes av mange for å være teksten hvor han faktisk langt på vei lyktes med å skape et filosofisk verk i tradisjonell forstand, arbeider mer systematisk og forfølger et emne på en lineær, etterfølgbar måte. Derfor passet det ham også med en omfattende innholdsfortegnelse til denne teksten.

DISCOVERY benytter slike innholdsfortegnelser av forfatteren som veiviser og støtte i arbeidet med å utvikle verktøy for semantiske søk. Samtidig bør den semantiske kodingen og slike verktøy mer generelt også kunne se helt bort fra forfatterens egen etikettering for heller å fange opp brukernes perspektiver og behov. Et avsnitt som for forfatteren har hørt til under emne x og derfor har preg av sjargong y, skal kunne få et etikett som relaterer det til emne z og til nyere språkbruk og begrepsdannelser. I den digitale tekstverden er denne type "oversettelse" lettvint og i grunnen mulig å utvide og revidere i det uendelige, og gjør verket vesentlig mer anvendelig og lettere å forstå. Dette er nyttig selv i verker som er godt strukturerte og til og med har innholdsfortegnelse, men enda mer nyttig er det når objektet er tekster med "criss-crossing" form og innhold – som for eksempel mange av Wittgensteins og Nietzsches tekster.

Tre prinsipper

Wittgensteinarkivets bidrag til DISCOVERY-prosjektet kjennetegnes av tre gjennomgående prinsipper:

1. separasjon av primærregistrering og presentasjon,
2. åpen kildekode og åpen tilgang, og
3. kollaborativ forskning.

De siste to har vært sentrale prinsipper i HyperNietzsche-prosjektet, men også Wittgensteinarkivets webutgivelser og spesielt prosjektene "Wittgenstein MS101 from September 1914" og "Tracing Wittgenstein" (se over) har praktisert åpen tilgang til tekster og programvare. I "Witt-

genstein MS101 from September 1914” er tekst- og faksimileutgaver gjort fritt tilgjengelige på Web, sammen med XML-transkripsjonene, DTD-filene, stylesheet-filer og de andre filene som lå til grunn for utgavene.

Separasjon av primærregistrering og presentasjon er et viktig prinsipp og en ”kjepphest” for Wittgensteinarkivet. Separasjon er grunnleggende ikke bare for den opprinnelige oppbyggingen av den maskinleselige versjonen og hele Wittgensteinarkivets edisjonsfilologiske arbeid på 1990-tallet, men også for all aktivitet i utgivelse og semantisk koding som pågår nå. Kort sagt går dette prinsippet ut på å skille spørsmål som er knyttet til registrering av manuskriptdata fra spørsmål om deres presentasjon for brukeren. Et eksempel på å ikke følge dette prinsippet ville vært å bruke MS Word til å merke tekst (som i original er overstrøket) med en overstrykingsfunksjon. Gjør man dette, blir det senere vanskelig å ”prosessere” overstrøket tekst på annen måte enn å vise den med MS Words overstrykingsfunksjon. Derimot tillater separasjonen av registreringsspørsmål og presentasjonsspørsmål – som skjer gjennom tekstkoding – å bruke informasjonen om at tekst er overstrøket på et mangfold av måter, og i tillegg åpner den for at brukeren selv kan avgjøre hvordan og om den brukes. Leseren finner mer om dette prinsippet og dets sammenheng med tekstkoding i Pichler & Bruvik (under utg.) hvor det gis uttrykk for synspunktet at det først med det digitale mediet og tekstkoding er blitt mulig å praktisere og fremme dette prinsippet, selv om det kan ha vært et ideal i edisjonsfilologien også mye tidligere.

Åpen kildekode (open source) og *åpen tilgang* (open access) har sin frammarsj i offentlig IT-politikk. Vitenskapelig utgivelse og publisering er et område hvor digitalisering og ny teknologi bidrar til å forandre praksis og vedtatte sannheter. Tradisjonelt publiseres forskernes artikler i tidsskrifter som tjener penger på å selge artikler i et marked. Fagbibliotekene er den viktigste salgskanalen for vitenskapelige artikler. Mange mener en slik ordning er fundamentalt gal: først skal det offentlige finansiere forskning, som så publiseres i forlag som definerer en høy pris. Deretter må det offentlige kjøpe tilbake forskningsresultatene gjennom bibliotekene (se Hillesund 2005). Stadig flere krever nå *åpen tilgang* til forskningsresultater fra offentlig støttet forskning og at forskere publisere

rer sine artikler gjennom såkalt åpen publisering som gir brukerne fri tilgang til innholdet fra de elektroniske online-arkivene.

Også EU er bekymret over at tilgjengeligheten til vitenskapelig publisert materiale er for dårlig, og for den dominerende rollen til de kommersielle vitenskapelige tidsskriftene og forlagene. I EUs eContentplus-program er et hovedanliggende å få til mer gjenbruk av vitenskapelig og kulturelt digitalt innhold i Europa. Det bør vurderes nye og mer egnede verdikjeder for forskningspublisering i kunnskapssamfunnet, mener EU-kommisjonen. Dette er viktig ikke bare for å utnytte mulighetene med ny teknologi helt generelt, men også for å legge til rette for en mer effektiv europeisk forskning og innovasjon.

DISCOVERY har forpliktet seg til å praktisere både åpen kildekode og åpen tilgang. Det betyr at koden til DISCOVERYS programvare vil være åpent tilgjengelig og vil kunne viderebrukes av alle, og at også tekstene og ressursene som bygges i prosjektet vil være åpent tilgjengelige og kan brukes av alle. Prosjektet er klar over at den åpne interaksjonen det legges opp til krever en velfungerende teknologi, men også høy kvalitetssikring av innholdet som gjøres tilgjengelig for alle. Før nytt materiale skal legges ut og før brukerredigering av allerede eksisterende materiale blir synlig for alle, må det gjennom et fagfelle-evaluerende system. Det vil også praktiseres flere "Web-rom" som kan ha ulike nivåer av kvalitetssikring. DISCOVERY tilstreber altså like gode rutiner for kvalitetssikring som i dag er brukt for publisering i kommersielle fag-tidsskrift.

Wittgensteinarkivet skal i DISCOVERY gjøre 5 000 sider Wittgenstein *Nachlass*-materiale fritt tilgjengelig på Web, både som faksimile og tekstutgivelse. Dette forutsetter selvfølgelig en overenskomst med de andre involverte rettighetshavere. I 2007 har Wittgensteinarkivet fått den nødvendige tillatelsen av Oxford University Press (OUP) og Trinity College Library, Cambridge. OUP og Wittgensteinarkivet hadde på 1990-tallet inngått kontrakt om publisering av *Nachlass*, som krever at OUP gir sin tillatelse til WABs bidrag i DISCOVERY-prosjektet. Det faktum at OUP er med på prosjektet viser at prosjektets utgangspunkt om forenligheten av kommersiell publisering og åpen tilgang ikke er galt. Trinity College Library har, etter at den siste opprinnelige testamentsforvalter av Wittgensteins *Nachlass*, Georg Henrik von Wright, døde i

2003, overtatt funksjonen til Wittgenstein "Trustees" og er nå institusjonen som forvalter rettighetene i og bruk av *Nachlass*. I tillegg har de enkelte bibliotekene og arkivene hvor *Nachlass*-materiale befinner seg rettigheter. Det anses som en stor suksess at det er mulig å gjøre 5 000 sider av Wittgensteins *Nachlass* fritt og åpent tilgjengelig på Web. Disse sidene vil innbefatte noen av de mest "populære" skriftene til Wittgenstein, som for eksempel "Big Typescript".

Kollaborativ forskning er viktig i DISCOVERY. I prosjektet bidrar både prosjekt-teamet og "eksterne" brukere til webressursen med metadata tagging, kommentarer, forklaringer og egne transkripsjoner. Nye bidrag vil ofte være komplementære til eller konkurrerende med materiale som er publisert på ressursen fra før. Dette fører gjerne til at tidligere bidrag blir rettet på, men uten at det opprinnelige bidrag vil forsvinne fra nettet siden referansestrukturen må ivaretas gjennom hele prosjektet (og utover). DISCOVERY blir altså i mindre grad en portal til faksimiler og utgivelser av filosofimateriale enn en integrert plattform for kollaborativ forskning omkring materialet. I tillegg arbeider DISCOVERY for en bedre integrering av individuell og kollaborativ forskning med undervisning. Det vil utvikles en virtuell infrastruktur som kan integrere vitenskapelig publisering, kollaborativ forskning og e-læring. Det tradisjonelle skillet mellom forskning på høyt nivå og "text book" læring aksepteres ikke i dette prosjektet.

For humanistisk forskning og studier synes nyttepotensialet ved slike prosjekter som DISCOVERY enormt. Men som for all slik innovasjon: det handler ikke først og fremst om teknologienes forløsende egenskaper, det handler mest om fagmiljøenes og enkeltindividenes evne til å spille med teknologien. Og deres evne til å få øye på, ta inn over seg og aktivt utnytte de mulighetsrom som er der. Internett og digitalisering endrer samfunnets virkemåte og samhandlingsmåter på fundamentale vis. Ikke minst endres rammebetingelsene for forskning og utdanning. I hvert fall blir det klart synlig at Wittgensteinarkivets arbeid og kanskje humanioras arbeid mer generelt foregår vesentlig i en internasjonal sammenheng, og i betydelig grad er et kollaborativt prosjekt som utføres i samarbeid mellom komplementære partnere fra flere land. På samme måte som Wittgensteinarkivet er en sentral aktør i disse prosjektene, ville det ikke vært mulig å drive Wittgensteinarkivet på en adekvat måte uten bi-

drag fra de utenlandske partnerne. Dette gjelder deres filosofisk-faglige, IT-faglige, så vel som deres finansielle bidrag. Her er det mange praktiske aspekter som kanskje kan si oss noe vesentlig om de utfordringer og muligheter digitale humaniora står ovenfor med hensyn til organisering av utvikling og vedlikehold av digitale ressurser, og deres videre bruk i forskning og læring.²

Alois Pichler er leder for Wittgensteinarkivet og forsker ved Avdeling for kultur-, språk- og informasjonsteknologi (AKSIS) ved Unifob/Universitetet i Bergen, Norge.

E-post: Alois.Pichler@aksis.uib.no

Webbplats: <http://teksttek.aksis.uib.no/people/14>>

Gjermund Lanestedt er seniorkonsulent i norsk Teleplan AS og norsk nasjonalt kontaktpunkt for EU-programmet eContentplus (2005-2008).

E-post: gjermund.lanestedt@teleplan.no

Webbplats: <http://www.digi.no/php/bransje.php?id=12387>>

Noter

1. Vi takker Hans Walter Gabler, Odd Einar Haugen og Kjell S. Johannessen for kommentarer til utkast av denne artikkelen.
2. DISCOVERY-prosjektet ble tidligere presentert i kort form i Lanestedt 2006.

Referanser

Alle URLer i denne artikkelen er kontrollert 22. august 2007.

BARBERA, MICHELE & RICCARDO GIOMI (2003). "The Pearl-Diver Model: The HyperNietzsche Data Model and its Caching System." *Jahrbuch für Computerphilologie* 5: 193-203.

BARBERA, MICHELE (2006). "The HyperLearning Project: Towards a Distributed and Semantically Structured e-research and e-learning Platform." *Literary and Linguistic Computing* 21.1: 77-82.

BIGGS, MICHAEL & CLAUS HUITFELDT, RED. (1997). "Philosophy and Electronic Publishing." *The Monist*. Interactive Issue 80.3: 348-367.

CRIPPS, PETER & ESPEN S. ORE (1997). "Elektronisk publisering av Wittgenstein's Nachlaß." *Human IT* 1.4: 47-54. <<http://www.hb.se/bhs/ith/4-97/pc-eso.htm>>

D'IORIO, PAOLO (2001). "Principes de l'HyperNietzsche." *Diogène* 196: 77-94.

EGGERT, PAUL (2005). "Text-encoding, Theories of the Text, and the 'Work-Site'." *Literary and Linguistic Computing* 20.4: 425-435.

HILLESUND, TERJE (2005). "Kommersiell istedenfor åpen publisering: en kritikk av det norske systemet for dokumentasjon av vitenskapelige publikasjoner." *Norsk medietidsskrift* 2: 155-165.

HRACHOVEC, HERBERT (2000). "Wittgenstein on line / on the line. " WAB: "Fragments". <http://wab.aksis.uib.no/wab_contrib-hh.page>

HUITFELDT, CLAUS (2001). "Utgivelsen av 'Wittgensteins Nachlass'." *Bok og skjerm*. Red. J.G. Jørgensen, Tone Modalsli, Espen S. Ore & Vigdis Ystad. Oslo: Fagbokforlaget. 115-133.

LANESTEDT, GJERMUND (2006). "I cyberinfrastrukturens tidsalder: europeiske

filosofimiljøer trekker veksler på semantisk metadata-tagging." *Computerworld* 28 (1. september): 36.

PICHLER, ALOIS (2002). "Encoding Wittgenstein: Some Remarks on Wittgenstein's Nachlass, the Bergen Electronic Edition, and Future Electronic Publishing and Networking." *TRANS: Internet-Zeitschrift für Kulturwissenschaften* 10.
<<http://www.inst.at/trans/10Nr/pichler10.htm>>

PICHLER, ALOIS & ODD EINAR HAUGEN (2005). "Fra kombinerte utgaver til dynamisk utgivelse: erfaringer fra edisjonsfilologisk arbeid med Wittgensteins filosofiske skrifter og nordiske middelaldertekster." *Læsemåder: Udgavetyper og målgrupper*. Red. Per Dahl, Johnny Kondrup & Karsten Kynde. København: C.A. Reitzels Forlag. 178-249.

PICHLER, ALOIS & TONE MERETE BRUVIK (under utg.). "Scholarly Digital Editions and the Principle of Separating Presentation from Registration." *Scholarly Digital Editions and Reading on Digital Support*. Eds. P. Régnier, C. Bélisle & D. Apollon.

ZÖLLNER-WEBER, AMÉLIE & ALOIS PICHLER (2007). "Utilizing OWL for Wittgenstein's Tractatus." *Philosophy of the Information Society. Papers of the 30th International Wittgenstein Symposium*. Eds. H. Hrachovec, A. Pichler & J. Wang. Kirchberg am Wechsel: ALWS. 248-250.