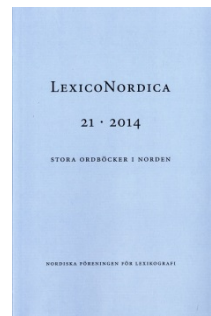


LexicoNordica

Titel: *Norsk Ordbok* i den digitale tidsalderen
Forfatter: Christian-Emil S. Ore & Åse Wetås
Kilde: LexicoNordica 21, 2014, s. 121-140
URL: <http://ojs.statsbiblioteket.dk/index.php/lexn/issue/archive>



© LexicoNordica og forfatterne

Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

Norsk Ordbok i den digitale tidsalderen

Christian-Emil S. Ore & Åse Wetås

A common challenge for the editing of historical and dialect dictionaries is the heterogeneity of the source material. *Norsk Ordbok* (NO) aims at providing a scholarly and exhaustive account of the vocabulary of Norwegian dialects from 1600 to the present and of the modern written standard *Nynorsk*, thus facing the heterogeneity in both dimensions. In 2002 a new and completely digital editorial system with a meta-index as a pivot part was implemented for NO. This index enables linking of each dictionary entry to its relevant available digital source material. The online dictionary has become an academic 'enhanced publication' with reproducible results. This well-defined interlinking of the digital material opens up for user-friendly interfaces and advanced search functions, illustrations, maps and publishing of the dictionary as Linked Open Data.

1. Innleiing

Denne artikkelen vil ta for seg dokumentasjonsverket *Norsk Ordbok* (NO) og ut frå både innhald og databasarkitektur drøfta korleis verket som digital ressurs kan formidlast på nye måtar og på nye publiseringsplattformer. Me vil dessutan sjå nærare på korleis ny funksjonalitet kan leggja til rette for at verket får auka bruksverdi også for nye brukargrupper.

Etter ei kortare innleiing om NO vil me sjå nærare på dokumentasjonsordboka som database. Struktureringa av informasjonen som ligg i det digitale kjeldematerialet og i dei ferdige digitale ordartiklane har gjort det enkelt å setja opp ein funksjonell nettversjon av verket. Artikkelen vil drøfta vidare utviklingsmoglegheit og utnyttingspotensiale for desse ressursane. Siste avsnitt inneheld oppsummering og konkluderande merknader.

2. *Norsk Ordbok* – bakgrunn og arbeidsform

2.1. Verket *Norsk Ordbok*

NO er eit dokumenterende ordboksverk som gjev ei vitenskapleg framstilling av det norske folkemålet – alle dei norske dialektane i by og bygd over heile landet – og det nynorske skriftmålet. Verket representerer ei vidareføring av Ivar Aasens leksikografiske verk (Aasen 1850, 1873), slik det m.a. er framstilt i Skard (1932:17ff.). NO byggjer på eit stort empirisk grunnlagsmateriale, og dette materialet er samla inn og systematisert gjennom heile utgjevingsperioden for verket.

Frå starten av var tanken å gje ut eit verk på 2–3 ordboksband, men redaksjonen såg snart at materialet og artikkellengda gjorde eit slikt omfang heilt urealistisk (jf. Hellevik 1966). Då første band kom ut, dekte dette alfabetbolken *a–doktrinær*, og grunnlagsmaterialet var i stadig auke. På opplauspsida av arbeidet med den trykte utgåva av verket nå i 2014, er den endelege fasiten at NO blir eit dokumentasjonsverk i tolv band. Kvart av banda inneheld om lag 25 000 ordartiklar, fordelte på 800 sider. I den moderne prosjektorganisasjonen NO 2014 har kvart band ei produksjonstid på 15 månader, og det omfattar om lag 35 årsverk fagleksikografisk arbeid.

I perioden 1966–2002 kom dei fire første banda av NO ut. Denne eldste delen av verket er ikkje digitalt tilgjengeleg, og det inneber at NO som digital ressurs førebels ikkje omfattar alfabetstrekket *a–h*.

På 1990-talet blei vesentlege delar av grunnlagsmaterialet til NO digitalisert gjennom det dåverande Dokumentasjonsprosjektet (sjå Ore & Kristiansen 1998). Då det norske Kulturdepartementet og Universitetet i Oslo ved tusenårsskiftet blei samde om at NO i 12 band skulle gjerast ferdig til feiringa av 200-årsjubileet for den norske Grunnlova i 2014, gjorde digitaliseringa av kjeldene

bak verket det mogleg å gå over til fulldigital ordbokproduksjon, der mange forskarar kunne arbeida inn i eitt felles elektronisk produksjonssystem.

Det tidsavgrensa prosjektet NO 2014 starta 1. juni 2002. I startfasen blei det lagt ned mykje arbeid i å utvikla ei elektronisk redigeringsplattform for ordboka. Arbeidet skjedde i nært samarbeid mellom Eining for Digital Dokumentasjon (EDD) og fagleksikografane i NO-redaksjonen. Resultatet er blitt ein virtuell fabrikk til bruk i systematisering, prosessering, tolking og sluttlagring av store mengder språkdata. I løpet av prosjektperioden er dette systemet blitt bygd ut med nye komponentar, m.a. med ein eigen modul for semantisk sortering (jf. Bakken & Grønvik 2008:31ff.). Verktøya er òg blitt lagde til rette for lagring av data i nye format. Den nyaste versjonen av NO 2014s kjeldeapplikasjon tillèt til dømes lagring av bilete og websider. Dette mogleggjer innhausting av veldokumenterte grunnlagsdata frå kjelder som ikkje er knytte til ordningar for stabil langtidslagring (jf. drøftinga av problemstillingar som gjeld tekstinnausting frå nettkjelder i Karlsen & Wetås 2010:523f.).

Parallelt med det digitale utviklingsarbeidet har NO 2014-organisasjonen utvikla redaksjonelle retningslinjer tilpassa den digitale produksjonen. Prosjektet har òg frå 2003 av bygd opp eit elektronisk tekstkorpus over nynorsk tekst frå 1860 til i dag. Dette korpuset er fullintegrrert i redigeringsverktøya og passerte tidleg i 2014 100 millionar ord.

Oppdraget frå Kulturdepartementet var opphavleg eit trykt ordboksverk i tolv band, medan ei nettutgåve av verket ikkje var med i dei originale planane. Med tida blei det meir og meir klårt at nettet i den digitale tidsalderen er ein viktigare formidlingskanal enn dei trykte bøkene. Kulturdepartementet har difor utvida «bestillinga» til å omfatta eit «gratis digitalt oppslagsverk på Internett». Kva dette inneber, er omtalt seinare i denne artikkelen.

Produksjonen av trykksatsen i NO 2014s redigeringsystem er

basert på XML-teknologi, og det var difor rett fram å publisera artiklane på nettet. I mars 2012 lanserte NO 2014 ei første nettutgåve av ordboksverket. I 2014 er prosjektorganisasjonen inne i sitt siste ordinære driftsår, og arbeidet med vidare tilrettelegging og raffinering av nettutgåva av ordboka og nettpublisering av det store grunnlagsmaterialet i databasane har høg prioritet. Ordboka blir publisert både som trykt tekst og som enkeltvis ordartiklar i nettutgåva. Band 12, som dekkjer alfabetstrekket *u–å*, skal etter planen ferdigstillast til desember 2014. Nettpubliseringa av artiklane i dette alfabetstrekket skjer samstundes med publiseringa av det trykte bandet.

2.2. Innhaldet i databasen og mellom permanente

I ferdig utgåve vil NO innehalda meir enn 300 000 store og små portrett av ord frå norsk folkemål og nynorsk skriftmål. I den elektroniske ordbasen av ferdige artiklar er alle ordartiklane lenkja opp til relevant grunnlagsmateriale. Dermed skal alle tolkingane redaktørane har gjort i arbeidet med ordartiklane i verket, kunna etterprøvast mot kjeldematerialet som ligg føre. Denne koplinga har vore med sidan starten av NO 2014-prosjektet i 2002. I dei seinare åra har slik utvida publisering («enhanced publication») blitt stadig meir utbreidd og metodar knytte til denne er etablerte som eit eige felt innanfor digital humaniora (sjå til dømes Ver-nooy-Gerritsen 2009).

I den leksikalske databasen er ein stor del av det norske ordtilfanget vitskapleg behandla og standardisert for første gong. Normeringa av ord frå folkemålet skjer ut frå ei grunnleggjande oppfatning om at alle orda i talt norsk kan ha eit skriftleg, nynorsk korrelat. Sjølv om delar av tilfanget er i geografisk avgrensa bruk, gjeld dokumentasjonen i NO i hovudsak det sentrale og mest brukte ordtilfanget i norsk. Intensjonen er at ordboksverket skal vera eit referanseverk for norsk språk i høve til både fonologi,

morfologi, semantikk og syntaks, slik det kjem til uttrykk gjennom både einskildord og (leksikaliserte) frasar.

3. *Norsk Ordbok* som database

3.1. Metaordboka – ein felles indeks og eit ordboks-univers

Som nemnt i punkt 2.1 blei vesentlege delar av grunnlagsmaterialet til NO digitalisert på 1990-talet gjennom det dåverande Dokumentasjonsprosjektet. I hovudsak dreidde dette seg om setelsamlinga med 3,2 millionar ordsetlar, men òg ei lang rekkje eldre og nyare ordsamlingar og ordbøker blei digitaliserte. Materialet var svært heterogent, og ikkje minst framstod det svært ueinsarta når det galdt tilhøvet til gjeldande eller tidlegare rettskrivingar. Oppslagsorda på setlane og i ordsamlingane var ofte ikkje i samsvar med noka gjeldande rettskriving. For å samordna eit så stort og mangfaldig materiale blei det laga ein felles indeks – Metaordboka (MO), der alle førekomstar av eit gjeve ord er samsorterte under grovnormerte oppslagsord (jf. Ore 2001:267ff.).

MO er meir enn ei felles søkjeliste. Artikkane her inneheld både ein oppslagsdel med informasjon om det normerte oppslagsordet (ordklasse, leddeling og rettskrivingsnorm) og lenkjer til ulike skriftrealiseringar og bruksdøme i dei enkelte deldatabasane som setelarkiv, eldre og nyare ordsamlingar og ordbøker. Ein artikkel i MO kan i prinsippet ha fleire oppslagsord som følgjer ulike rettskrivingsnormer. Dermed kan den til dømes utgjera ein felles indeks over bokmål og nynorsk. Den nåverande MO har oppslagsord på nynorsk etter 1938-rettskrivinga, men me planlegg å leggja til oppslagsord i moderne nynorsk rettskriving.¹

1 Det norske Språkrådet har i eit notat til Kulturdepartementet om norsk ordbokpolitikk (Språkrådet 2014) tilrådd at *Norsk Ordboks* nettutgåve

MO omfattar alt frå gamle ordsamlingar frå 1600-talet til moderne dialektsamlingar og ordbøker. I tillegg kan ein også lagra konkordansar frå det nynorske tekstkorpuset direkte i systemet. Metaordboka har vist seg å vera eit sær sars verdfullt verkty i redigeringa av NO, men er òg i seg sjølv ei verdfull samling (jf. her òg konklusjonane om denne ressursen i evalueringa av dei vitenskaplege språksamlingane ved Universitetet i Oslo av 21.3.14 (Kristoffersen, Ebeling, Vangsnæs & Wetås 2014)). Både dei eldre og nyare ordsamlingane er lagra i original form så langt som råd. Dei uttrykkjer språksynet til forfattaren og tida dei blei skrivne i, i tillegg til å gje lingvistisk og leksikografisk informasjon. Dei er òg historiske tidsvitne og utgjer ei viktig kjelde til kulturhistorisk dokumentasjon.

Lenkjene internt i MO og mellom Metaordboka og NO går begge vegar. Eit oppslagsord i ei dialektordbok gjev på denne måten tilgjenge til alle dei andre ressursane i systemet, inklusive den tilhøyrande artikkelen i NO. Dette gjev dialektbrukarane ei eineståande moglegheit til å sjå eigen dialekt i ein større kontekst.

3.2. *Norsk Ordbok* frå tekst til tre – analyse og resultat

I dag blir dei fleste ordbøker redigerte ved hjelp av XML-teknologi. Det vil seia at ein definerer ein temmeleg fast hierarkisk trestruktur som artiklane må følgja. Kor fast denne strukturen skal vera, varierer frå ordbok til ordbok, men XML-teknologien gjer det enkelt å produsera trykksats og publisera ordboksartiklane på nett. For eldre ordboksverk som er retrodigitaliserte, er det ikkje alltid mogleg å finna eit stramt XML-skjema. Det er nok ikkje utan grunn at tilrådingane frå Text Encoding Initiative (TEI P5) har eit eige XML-element «entryFree» for tekst-

bør ligga føre i nynorsk etter den vedtekte 2012-norma. For å leggja til rette for brei bruk av grunnlagsmaterialet, er det viktig at også MO får oppslagsord etter gjeldande norm.

koding av ordbøker, der delelementa kan komma i vilkårleg rekkjefølgje.

Dei to første banda av NO blei redigerte manuelt (skrivemaskin og tradisjonell sats). Manuskriptteksten til band 3, 4 og halve band 5 blei som ein del av redigeringsarbeidet skrivne inn i eit tekstbehandlingsprogram og sett på grunnlag av eit enkelt kodesystem. Tekstflyten i manuskripta følgjer den tradisjonelle tanken om at ordboka er ein tekst, og ikkje nokon database eller trestruktur. Å tilpassa teksten i desse banda til eit moderne redigeringsystem krev difor same analysen som for andre eldre ordboksverk.

Då redigeringsystemet for NO 2014 skulle lagast i 2002, blei teksten i dei eksisterande, eldre banda djupanalysert. Tanken var å identifisera og vidareføra den beste praksisen i systematisk form (jf. òg Grønvik 2005:161), og prosjektet var ikkje bunde til å laga eit system som kunne handtera den eldre teksten fullt ut. Mellom anna såg ein at redaktørane mange gonger i dei manuelt utforma artiklane hadde lagt inn tilvisingar til ordsamansetningar som høyrde til lengre ute i alfabetet. I eit komplett verk og i ei elektronisk ordbok må dette gjerast gjennom elektroniske krysstilvisingar.

I analyseprosessen valda òg faste ordsamband og frasar ein god del bry. I dei fleste ordbøker blir slike handsama som bruksdøme med forklåring. Eit døme er frasen *panser og plate*. I NO 2014-løysinga har denne fått ein eigen artikkel i artikkelformatet «underartikkel», og med lenkje til hovudartikkelen som fleirordsfrasen høyrer saman med. I dette tilfellet kunne ein tenkja seg lenkje både til ordartikkelen **panser** og artikkelen **plate** (som begge er med i NO), men ut frå dei vedtekne redaksjonsreglane i NO 2014s redigeringsmanual (RH), skal underartikkelen **panser og plate** lenkjast opp mot artikkelen **panser**.² På trykk er slike underartiklar innbakte i teksten til hovudartikkelen. I den elektroniske ordboka er dei moglege å søkja opp både som ein del av hovudartikkelen

² Nedfelt i RH kap. 4.3.8.4.1–2.

og som eigne, sjølvstendige artiklar med lenkje tilbake til hovudartikkelen, sjå figur 1. Alle andre tilvisingar, inklusive fullsynonym, er òg uttrykte gjennom lenkjer mellom artiklar.

The screenshot shows the interface of the Norwegian Dictionary (Norsk Ordbok 2014). At the top, it says "Norsk Ordbok 2014" and "Ordbok over det norske folkemålet og det nynorske skriftmålet". Below this is a search bar with the text "panser og plate" and a "Søk" button. To the right of the search bar are links for "Bibliografi" and "Avansert søk". Below the search bar are navigation links: "FØRRE", "TREFFLISTE", "ARTIKKEL", and "NESTE". The main content area shows the search results for "panser og plate". It lists two entries:

1. full rustning (Litterær kjelde: Fedrah 1878)
krigarar til hest vart kledde i panser og plate (Litterær kjelde: Stokk.N1 72)
2. OVERFORD TYDING
(dei har) stengt humaniteten sin inne bak panser og plater (Litterær kjelde: Borg.NKB 396)

Figur 1: Underartikkelen *panser og plate* i NOs nettutgåve.

I tillegg til tilvisingane mellom artiklane er alle brukseksempel kjeldefeste og kopla til ein bibliografibase med bibliografiske opplysningar uttrykte i formatet til biblioteksstandarden MARC (Machine-Readable Cataloging). Geografisk heimfesting er kopla til ein database over den norske heradsinndelinga frå 1947. Kvar heimfesting viser anten til eit herad eller eit større område som er ei gruppe av slike. Sidan polygonane som definerer grensene til herada er georefererte (koordinatfeste), er det mogleg å visa heimfestingane på kart, som vist i figur 2.

smøye-stol substantiv, hankjønn

1. hempe el lydkje på reim til å smøya (1a), stikka reimenden (på belte) inn i; smygstol (Heimfestingar: heller vanl, Rom R, Fet, Odal, Sol, Stange, Ringsaker, Åmot, Solli, Dovre, Heidal, VGausdal, NLand, ØToten, Ri, Lier, YSandavær, Tel R, Orangedal, Gjerstad, Hjelmeland, Suldal, Kvinherad, Tysnes, Strandebarm, Vossestrand, Lindås, Eid i N5, Haram, Roms, Ndm sumst, nFjells **Kårt**)
du må trø reima igjennom smøvestolen (Litterær kjelde: SuldalsmS.Sandv.)

2. ring til å smøya (1a), stikka eitkvart gjennom, (serf:) ring, bøyle på hestesele som taumane går gjennom; smettestol (2), taumstol (Heimfestingar: Rom R, Eidsvoll, Odal R, Heidal, Hall A2, Tel R, Kvitesæid, Ryf R-4, Sunndal, Tri **Kårt**; Litterær kjelde: NFL11Ød 26); òg nytta som stadnamn (Litterær kjelde: NI II,170)

3. stol av korgfletting (Heimfesting: Lærdal **Kårt**); jfr smøya (1b)

Stockholm
Hels
Statens kartverk

Figur 2: Ordartikkelen **smøvestol** i NOs nettutgåve. Kartet viser utbreiing av definisjonen i tydingsnummer 1.

Redigeringsystemet til NO er bygd opp av ulike modular som er knytte saman ved hjelp av referansar. Det har fleire gonger blitt sett spørjeteikn ved kvifor relasjonsdatabasesystem er valt i staden for XML. Svaret på dette er at NO 2014s redigeringsssystem nyttar begge typane. Grunnen er at relasjon databasar er best til noko av funksjonaliteten den store vitskaplege ordboka må ha, medan XML-teknologi er betre til andre delar av denne. Kryssreferansar og dataintegritet blir best teke hand om av relasjonsteknologien. Komplekse og raske søk er enklast å få til med ein relasjonsdatabase. XML-teknologien er særskild god til å lagra tekstleg informasjon og manipulera han. Det er ikkje uvanleg at informasjonssystem er baserte på begge teknologiane (sjå også drøftinga av dette i Ore & Tvedt 2006:318f.).

I NO 2014-systemet er ordboksteksten under redigering lagra i ein relasjonsdatabase, medan artiklane i ulike stadium i redigeringsprosessen (frå ny til ferdig) er lagra i ein logg som XML-dokument. I prinsippet kunne me utan problem erstatta redigeringsgrensesnittet vårt med ein såkalla XML-editor. Men ein slepp uansett ikkje unna programmodular som held orden på og sjekkar konsistensen på peikarane mellom artiklar, til andre delbasar som inneheld informasjon om stad, bibliografi, personar, eller som omfattar meir administrative metadata. Då trengst noko meir, dvs. tabellar og relasjonar. Redigeringsgrensesnittet for artiklane i NO er gjennom åra finslipt slik at det skal vera lett å læra opp nye redaktørar. I tillegg skal systemet støtta opp under rask produksjon og sørgja for at sluttproduktet er konsistent og i samsvar med grunnlagsempirien.

Det viktigaste skiljet er om ein vel å skriva ordartiklane på tradisjonell måte som fritekst utan formell struktur, eller om ein ser nytte i å organisera innhaldet som ein gjennomtenkt formell struktur. Debatten om XML kontra relasjonsdatabasar er etter vårt syn lite relevant for fagleksikografer. Avgjerande for ei ordbok sin systematikk og akribi er ein godt gjennomtenkt datamodell og struktur, kombinert med solid fagleg arbeid.

4. Nettutgåva av ordboksverket

Nettutgåva av NO er fritt tilgjengeleg og omfattar dei delane av ordboksverket som er produserte på ei digital plattform. Dette gjeld alfabetstrekke frå og med *i*. Det redaksjonelle produksjonssystemet som blei utvikla frå og med 2002, blei laga for plattform-uavhengig publisering. Dette medførte at arbeidet med å laga nettpublikasjonen NO var ein relativt liten og enkel jobb (sjå Ore 2012:492f., jf. òg Ore & Tvedt 2006:319f.). Den store faglege vinsten i å gje ut ordboka også som nettpublikasjon, ligg i at dette gjev oss høve til å ekspandera informasjon, slik det i NO på nett er gjort m.a. med kjelder, heimfestingar, etymologiopplysningar og ordklasseinformasjon. Det gjev vidare høve til å publisera visse kategoriar som tilvalsstoff for lesaren, slik at navigeringa i artikkelen blir enklare. I NO på nett er opplysningar om etymologi, uttale, alternative skriftformer og målførevariasjon lagde inn som informasjonskategoriar ordbokbrukaren må navigera seg inn i ved bruk av spesialutvikla ikon.

Nettutgåva har òg gjort det mogleg å leggja inn tilleggsfunksjonalitet som forklårande rettleiingstekst og aktive lenkjer på alle punkt der ordartiklar er kopla elektronisk saman på artikkel- eller tydingsnivå. I tillegg er kartvisningar frå 2014 integrerte for alle artiklar som har opplysningar om målførevariasjon eller geografisk avgrensa utbreiing for ein gjeven definisjon. Søkjefunksjonaliteten gjer det mogleg å søkja både på normalartiklar og på fleirordsfrasar i underartiklar (typen *slå an tonen, stolpar opp og ned, på naud og neppe* osv., jf. punkt 3.2).

Nettutgåva omfattar ikkje berre den teksten ein kan lesa i den trykte utgåva. Den skal vera eit digitalt oppslagsverk som omfattar opplysningane ein kan nå gjennom MO-koplingane, bibliografi- og kartdatabasane og dei store underliggjande samlingane av grunnlagsdokumentasjon. For den framtidige forvaltninga og utviklinga av NO er dette eit svært viktig poeng.

5. Utviklingsmoglegheiter og utnyttingspotensiale

5.1. Ny informasjon og nye plattformer

Den gjennomarbeidde datamodellen som ligg bak redigeringsprogrammet til Norsk Ordbok 2014 gjer det enkelt å kopla ny informasjon og nye funksjonar saman med dei eksisterande informasjonskategoriane. Det er mogleg å supplera ordartiklane med lyd, bilete og video, og det ligg eit stort potensiale både for forskning og formidling i å kopla saman den digitale ordbokressursen med andre system (til dømes Artsdatabanken og ensyklopediske ressursar som *allkunne.no* og *Store norske leksikon*). Ordboka i fullelektronisk format vil det òg vera mogleg å kopla saman med andre ordbøker i Norden. Her kan ein tenkja seg å bruka Linked Open Data-teknologi og såkalla semantiske teknologiar som RDF (Resource Description Framework), til eit slikt felles ordboksnett i Norden. Moglegheitene er legio, så lenge ein har veldefinerte modellar for ordboksdata.

Utviklinga av nye publiseringssplattformer bør også gjera at ordboksressursane blir lagde til rette for slike nye format. Særleg aktuelt vil det vera å laga ein versjon tilpassa smarttelefonar.

5.2. Språkvitskapleg utnyttingspotensiale

NOs digitale ordbase er utforma slik at han gjev eit svært godt utgangspunkt for utarbeiding av ordnett. Den definerte artikkelkategorien «underartikkel», som er til bruk på alle leksikaliserte fleirordsfrasar i NO-basen, gjer det òg mogleg å ekstrahera desse til oppbygging av eit separat konstruktikon (jf. Lyngfelt & Sköldberg 2013). I ei leksikalsk basert framstilling som ei ordbok vil dei minst generelle konstruksjonane vera dei mest tilgjengelege og dei som i størst omfang er skildra i egne artikkelstrukturar, medan

konstruksjonsmønster med større generalitet i mindre grad er omtalte, og kanskje berre kjem til uttrykk som bruksdøme (jf. til dømes Berg-Olsen 2014).

Organiseringa av informasjonen i ordbasen, med aktive digitale koplingar mellom ordartiklar, gjer det også mogleg å ekstrahe- ra taksonomiske nettverk for ulike fagfelt. Til dømes er folkelege artsnemningar for dyr og plantar alltid kopla saman med vit- skaplege artsnemningar, og dei vitenskaplege nemningane ligg inne i basen med søkbare bruksmerkingar (bot. og zool). Dette gjer basen verdfull for felt som til dømes etnobiologisk gransking (jf. Wetås 2013). Liknande utnyttingspotensiale finst på ei rekkje andre felt, og desse vil vera tilgjengelege innanfor rammene av ordboka som digital ressurs.

6. Oppsummering og konkluderande merknader

6.1. Det beste av gammalt og nytt

Å gje ut verket NO parallelt som trykt publikasjon og som nettpublikasjon kan framstå som unødvendig dobbeltarbeid. Me har i denne artikkelen vist korleis dei redaksjonelle redigeringsverktya er utvikla slik at resultatata kan publiserast på fleire plattformer utan ekstra tilrettelegging. I praksis viser det seg dessutan at den fleirmediale publiseringa hentar ut det beste av den gamle og den nye verda. Fordi NO 2014s oppdrag er å levera eit ordboksverk i 12 band à 800 sider trykt tekst til ein gjeven dato, må organisasjonen heile tida følgja strenge og ufråvikelege krav til omfangsdimensjonering og tidsbruk. Resultatet er informasjonstette og stringente artiklar.

Publiseringa på digitale plattformer gjev høve til å laga utvi- da og meir brukarvennelege artikkelvisningar, til å tilby brukarane

aktive, klikkbare lenkjer mellom artiklar og til å kopla saman den ferdige ordboksressursen med andre digitale ressursar. Med ein ferdig ordbase som dekkjer alfabetstrekket frå *a* til *å*, og der heile tekstmassen er utforma etter dei same redaksjonelle retningslinene, vil det opna seg nye moglegheiter for å ekstrahera delmengder av artiklane til annan bruk. Denne kan vera geografisk eller tematisk avgrensa, eller òg til dømes avgrensa av bruksfrekvens, til bruk i ordboksprodukt som skal dekkja kjerneordforrådet.

Den store ordboka på digital plattform har på denne måten gjeve tre viktige resultat: For det første er store mengder språkdokumentasjon gjort digitalt tilgjengeleg både for forskarsamfunnet i inn- og utland og for ålmenta. Dette gjeld både dei ferdige leksikografiske produkta og dei store underliggjande relasjonsdatabasane med digitalt tilgjengelege grunnlagskjelder.

For det andre har EDD og NO 2014 utvikla leksikografiske forskingsverktøy som møter behovet for lett etterprøvar dokumentasjon av språk. Desse verktøya har eit stort utnyttingspotensiale både for norsk og for andre språk med ulik grad av standardisering, og dei er dessutan sette opp slik at dei kan handtera diakron og synkron heterogenitet (jf. òg Grønvik & Ore 2013:244f.).

For det tredje har overgangen til digitale publiseringplattformer for den store ordboka gjort NO til ein ressurs som stadig kan byggjast ut og som i tillegg kan koplast saman med andre digitale ressursar. Ein kan her nemna ressursane publiserte på ordnet.dk (særleg DDO og ODS) og ei framtidig digital utgåve av SAOB. Desse inneheld relevant informasjon om nærskylde ord i dei skandinaviske språka. Sist men ikkje minst har ein den framtidige vitenskaplege ordboka over norsk bokmål (jf. Vonen 2014). Det er særskild viktig at ein planlegg eit dokumenterande verk over bokmålet slik at det blir lett å kopla denne ressursen opp mot *Norsk Ordbok*. Resultatet vil bli ei best mogleg felles skildring av dei to norske skriftspråka.

6.2. Sluttord om store ordbøker som digital ressurs

Fulldigital ordboksproduksjon bidreg vesentleg til ei systematisk og stringent behandling av ordmaterialet. I tillegg gjev den digitale produksjonen store vinstar i form av høve til å behandla store mengder data. For NOs del har omlegginga til fulldigital produksjon gjort at redaksjonen kan ha eit høgt arbeidstempo. Dette skuldast ikkje minst at mange tidkrevjande rutineoppgåver nå blir løyste av produksjonssystemet. Eit enkelt døme er måten litteraturreferansar, heimfestingar, grammatiske kategoriar, bruksmerkingar osv. er gjorde tilgjengelege som val frå faste menyar. Det sparar arbeid i redigeringsfasen, det sparar arbeid i korrekturfasen og det gjer det mogleg å enkelt ekspandera referansar og kategori-markeringar i digitale visningsformat.

Store og dokumenterande ordboksverk er av natur komplekse og informasjonstette. Dette gjer fleirbrukspotensialet deira som digitale ressursar endå større.

Litteratur

Ordbøker

DDO (2003–05) = *Den Danske Ordbog*. København: Det Danske Sprog- og Litteraturselskab og Gyldendal. <ordnet.dk/ddo> (august 2014).

NO (1966–) = *Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet*. Oslo: Det Norske Samlaget. Nettutgåve: <no2014.uio.no> (juni 2014).

ODS (1918–56) = *Ordbog over det danske Sprog*. København: Det Danske Sprog- og Litteraturselskab og Gyldendal. <ordnet.dk/ods> (august 2014).

- SAOB (1898–) = *Ordbok över svenska språket utgiven av Svenska Akademien (Svenska Akademiens Ordbok)*. Lund: Gleerups förlag.
- Aasen, Ivar (1850): *Ordbog over det norske Folkemaal*. Kristiania: Carl C. Werner & Comp.
- Aasen, Ivar (1873): *Norsk Ordbog*. Christiania: P.T. Mallings Boghandel.

Annan litteratur

- allkunne.no. Nynorsk digitalt oppslagsverk. <allkunne.no> (juni 2014).
- Artsdatabanken. <www.artsdatabanken.no> (august 2014).
- Bakken, Kristin & Oddrun Grønvik (2008): Materialsortering på digital plattform. Eit steg mot høgare dataintegritet i den vitenskaplege leksikografien. I: Asta Svavarsdóttir, Guðrún Kvaran, Gunnlaugur Ingólfsson & Jón Hilmar Jónsson (red.): *Nordiske Studier i Leksikografi 9. Rapport fra konference om leksikografi i Norden. Akureyri 22.–26. maj 2007*. Reykjavík 2008, 31–42.
- Berg-Olsen, Sturla (2014): Skjematiske fleirordskonstruksjonar i *Norsk Ordbok*. I: Ruth Vatvedt Fjeld & Marit Hovdenak (red.): *Nordiske studier i leksikografi 12. Rapport fra Konferanse om leksikografi i Norden. Oslo 13.–16. august 2013*. Oslo 2014, 80–94.
- Grønvik, Oddrun (2005): Redigeringsprogrammet for Norsk Ordbok 2014. I skjæringspunktet mellom menneskeleg skjønn og automatisering. I: Ruth Vatvedt Fjeld & Dagfinn Worren (red.): *Nordiske studiar i leksikografi 7. Rapport frå Konferanse om leksikografi i Norden. Volda 20.–24. mai 2003*. Oslo 2005, 156–165.
- Grønvik, Oddrun & Christian-Emil Smith Ore (2013): What should the electronic dictionary do for you – and how? I: Iztok Kosem et al. (eds.): *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference, 17–19 October 2013, Tallinn, Estonia*. Tallinn: Eesti Keele

- Instituut, 243–260. <eki.ee/elex2013/conf-proceedings/> (juni 2014).
- Hellevik, Alf (1966): *Innleiing*. I: Olav T. Beito et al. (red.): *Norsk Ordbok*. Band 1: *a–doktrinær*. Oslo: Det Norske Samlaget, I–XVI.
- Karlsen, Knut E. & Åse Wetås (2010): Internettpublikasjonar som kjelde til dokumentasjonsordbøker. Status og bruk. I: Harry Lönnroth & Kristina Nikula (red.): *Nordiska studier i lexikografi 10. Rapport frå Konferensen om lexikografi i Norden. Tammerfors 3–5 juni 2009*. Tammerfors 2010, 522–529.
- Kristoffersen, Kristian Emil, Signe Oksefjell Ebeling, Øystein Vangsnes & Åse Wetås (2014): Den forskings- og undervisningsmessige verdien av eit utval samlingsressursar ved ILN. Rapport frå ei arbeidsgruppe nedsett av instituttleiar ved ILN hausten 2013. <<http://bit.ly/1vdjAFS>> (juni 2014).
- Lyngfelt, Benjamin & Emma Sköldberg (2013): Lexikon och konstruktikon – ett konstruktionsgrammatiskt perspektiv på lexikografi. I: *LexicoNordica* 20, 75–92.
- MARC Machine-Readable Cataloging. <www.loc.gov/marc/umb/> (juni 2014).
- Ore, Christian-Emil (2001): Metaordboken – et rammeverk for Norsk Ordbok? I: Martin Gellerstam, Kristinn Jóhannesson, Bo Ralph & Lena Rogström (red.): *Nordiska studier i lexikografi 5. Rapport frå Konferens om lexikografi i Norden. Göteborg 27–29 maj 1999*. Göteborg 2001, 250–270.
- Ore, Christian-Emil (2012): Nettordbøker og Norsk Ordbok – hvordan etablere en vitenskapelig nettordbok. I: Birgit Eaker, Lennart Larsson & Anki Mattisson (red.): *Nordiska studier i lexikografi 11. Rapport frå Konferensen om lexikografi i Norden. Lund 24–27 maj 2011*. Lund 2012, 488–499.
- Ore, Christian-Emil & Nina Kristiansen (1998): *Sluttrapport frå Dokumentasjonsprosjektet*. <www.dokpro.uio.no/sluttrapp.pdf> (april 2014).

- Ore, Christian-Emil & Lars Jørgen Tvedt (2006): Digital sats eller digital satsing. I: Henrik Lorentzen & Lars Trap-Jensen (red.): *Nordiske Studier i Leksikografi 8. Rapport fra Konferanse om Leksikografi i Norden. Sønderborg 24.–28. maj 2005*. København 2006, 315–322.
- RDF = Resource Definition Framework. <www.w3.org/RDF/> (april 2014).
- RH = *Redigeringshandbok for Norsk Ordbok 2014*. <<http://bit.ly/109tCa1>> (juni 2014).
- Skard, Sigmund (1932): *Norsk Ordbok. Historie – plan – arbeidskipnad*. Oslo: Det Norske Samlaget.
- Språkrådet (2014): Norsk ordbokpolitikk. Notat fra Språkrådet til Kulturdepartementet 3. mars 2014. <<http://bit.ly/OBoRYU>> (april 2014).
- Store norske leksikon*. <snl.no> (juni 2014).
- TEI P5: *Guidelines for Electronic Text Encoding and Interchange*. TEI Consortium, Charlottesville, Virginia, USA 2014. <www.tei-c.org/release/doc/tei-p5-doc/en/Guidelines.pdf> (juni 2014).
- Vernooy-Gerritsen, Marjan (2009): *Emerging standards for enhanced publications and repository technology: Survey on technology*. Amsterdam: Amsterdam University Press.
- Vonen, Arnfinn Muruvik (2014): Behovet for store ordbøker i en samlet norsk ordbokpolitikk. I: *LexicoNordica 21* (dette nummeret).
- Wetås, Åse (2013): Plantenemningar i Norsk Ordbok. I: Lars-Erik Edlund, Inger Larsson & Staffan Nyström (red.): *Växter och växtamn – ett möte mellan botanik och språkvetenskap*. Acta Academiae Regiae Gustavi Adolphi CXXIII. Uppsala: Kungl. Gustaf Adolfs Akademien för svensk folkkultur, 165–174.

Christian-Emil S. Ore
førsteamanuensis, leiar for Eining
for Digital Dokumentasjon,
dr.scient.
Institutt for lingvistiske og nordiske
studium
Universitetet i Oslo
Postboks 1123 Blindern
NO-0317 Oslo
c.e.s.ore@iln.uio.no

Åse Wetås
prosjektdirektør NO 2014, dr.art.
Institutt for lingvistiske og nordiske
studium
Universitetet i Oslo
Postboks 1021 Blindern
NO-0315 Oslo
ase.wetas@iln.uio.no

