

Sluttrappport: «Tidsskriftsartikler kompatible med talesyntese», ref. 2021/132

Innhold

Prosjektbeskrivelse.....	1
Prosjektforløp	2
Første fase: Filer spesielt tilrettelagt for talesyntese.....	3
Andre fase: Universelt utformet EPUB	4
Videreføring og anbefalinger	5
Forarbeid	5
Rutiner	6
Oppsummering.....	6

Prosjektbeskrivelse

Nasjonalbiblioteket bevilget midler til en 80%-stilling over en periode på 20 måneder til prosjektet “Tidsskriftsartikler kompatible med talesyntese” (ref. 2021/132). Dette prosjektet hadde som formål å forbedre universell utforming av åpent tilgjengelige tidsskrifter publisert av UiTs publiseringstjeneste [Septentrio Academic Publishing](#), ved å gjøre dem mer kompatible med talesyntese, det vil si automatisk opplesning av tekst. Hovedansvaret for prosjektet lå hos Universitetsbiblioteket ved UiT (UBT). Norsk lyd- og blindeskriftbibliotek (NLB), som har bred erfaring med talesyntese, stilte som rådgivende part. Prosjektmedarbeideren hadde arbeidssted ved UBT, med arbeidskontrakt fom 15. november 2021 tom 14. juli 2023.

UBT drifter publiseringstjenesten Septentrio Academic Publishing, som utgir både fagtidsskrift og tidsskriftene som fokuserer på allmennrettet formidling. Alle tidsskrift er tilknyttet UiT Norges arktiske

universitet og er åpent tilgjengelige. Septentrio drifter en installasjon av Open Journal Systems (OJS) og hjelper tidsskriftene å holde god teknisk standard. I sammenheng med arbeidet om forbedring av universell utforming, ønsket UBT å undersøke mulighetene for å gjøre artikler kompatible med talesyntese. Dette ville gjøre fagtidsskriftenes innhold mer tilgjengelig for studenter og ansatte med synshemming, og for de som av andre årsaker drar nytte eller glede av automatisk opplesning.

Artiklene i Septentrio har tidligere bare vært publisert i PDF-format. Prosjektet har hatt som mål å utvikle en enklest mulig arbeidsflyt, inkludert dokumentmaler, for konversjon av innsendte manuskripter fra Word-format til et format kompatible med talesyntese. Som del av et annet tiltak, prøver UBT nå ut [XML](#) i tidsskriftet [Ottar](#), som kan gi flere artikkelformater (JATS XML, EPUB og HTML, eller XHTML). Talesyntese-prosjektmedarbeideren fant det som mest hensiktsmessig å koble talesyntese-relatert arbeidsflyt til XML-flyten og konkluderte med EPUB som det mest passende formatet. EPUB støttes av et bredt spekter av talesynteseprogramvare, inkludert NLB-appen, [Lydhør](#). I tillegg er EPUB et skalerbart format, dvs at det kan tilpasse seg ulike skjermstørrelser og oppløsninger for å optimalisere leseropplevelsen, noe som oppfyller kravene til universell utforming.

Prosjektet har eksperimentert med artikler i tre av Septentrios tidsskrifter ([Målbryting](#), [Nordic Perspectives on Open Science](#) og [Ottar](#)). Ulike løsninger har blitt testet på en intern server. Vi utviklet til slutt en arbeidsflyt i samarbeid med Ottar-prosjektet ved UBT. Alle programverktøy som benyttes i arbeidsflyten er åpen kildekode: OJS, [Pandoc](#) (brukes for dokumentkonvertering) og [epubJSViewer](#) (brukes i OJS for å vise EPUB-filer).

Prosjektforløp

Hovedlinjen i prosjektet kan deles inn i to faser. I den første fasen fokuserte vi på å tilrettelegge dokumentene for formater spesifikt tilknyttet til talesyntese, som SSML ([Speech Synthesis Markup Language](#)), CSS [Speech Module](#) og PL ([Pronunciation Lexicon](#)). Som beskrevet i den detaljerte oversikten av denne fasen (se seksjonen under), karakteriseres arbeidsflyt i denne fasen av betydelig mengde etterarbeid – derfor besluttet vi å prøve ut andre løsninger. I den andre fasen av prosjektet forsøkte vi i stedet automatisk konvertering fra Word til en standard EPUB. Dette ble gjort ved å basere oss på standardiserte Word-maler, som ble tilrettelagt for tilgjengelighet. Hensikten med dette var å sikre at dokumentene ble produsert fra bunnen av på en måte som var enkel nok til at forfattere og redaktører ikke behøvde kontinuerlig støtte fra Septentrio-teamet. Vi søkte å minimere behovet for etterarbeid, og sikre at dokumentene var i samsvar med Plan S, kravene til universell utforming og talesyntese.

Første fase: Filer spesielt tilrettelagt for talesyntese

I vårt første forsøk prøvde vi å konvertere Word-dokumentene til Markdown, for deretter å konvertere dem fra Markdown til XML med SSML-formatering. SSML er et XML-basert språk som gir utvidet kontroll over talesyntese, som uttale, tonehøyde, hastighet og volum. Imidlertid viste det seg at denne tilnærmingen var upraktisk da den ikke uten videre var forenlig med tidsskriftenes normale dokumentmal og krevde betydelig etterarbeid.

Vi prøvde oss deretter med en metode der vi konverterte fra Word til HTML, og brukte et stilark som inneholdt CSS Speech Module og Pronunciation Libraries (PL). CSS Speech Module er en utvidelse av Cascading Style Sheets (CSS), som gir kontroll over taleutgaven av en nettside, mens PL er biblioteker som kan brukes for å gi veiledning til korrekt uttale av ord. Denne tilnærmingen viste seg imidlertid også å ha sine begrensninger. Å utvikle egne PL-biblioteker ville ikke være realistisk gitt det omfattende arbeidet som ville kreves, og det relativt begrensede behovet. I tillegg, er ikke HTML lett å laste ned som en enkelt fil, og dette kan ekskludere brukere som foretrekker å bruke nedlastede filer. Vi bestemte oss derfor for å prøve ut samme stilark-løsning i EPUB-format.

EPUB er en åpen e-bokstandard forvaltet av W3C (World Wide Web Consortium). EPUB-formatet er allerede utbredt og har støtte for mange av de samme teknologiene som brukes i moderne webdesign, inkludert HTML, CSS, og JavaScript, samt mange tilgjengelighetsfunksjoner. Formatet rommer også nødvendige akademiske metadata, og samler all grafikk, tabeller og andre tillegg i en enkelt fil. EPUB-filene kan lages etter en felles mal og med et felles stilark (CSS), som i sin tur kan utformes i tråd med talemøduler og generelle prinsipper for universell utforming.

Tilgjengelighetshensyn inkluderer alternativ tekst til bilder, skalerbar størrelse med paginering, høy kontrast, valgfri skrifttype, og semantisk rekkefølge for å kunne navigere gjennom dokumentet ved å bruke "tab"-tasten, for eksempel for å hoppe mellom overskrifter.

En fordel med CSS-stilark er at de kan tilpasses én gang, og deretter brukes på tvers av mange dokumenter, noe som eliminerer behovet for individuell tilpasning av hvert dokument. Men, slik det er i dag, er det svært få talesyntese-programmer som støtter CSS Speech Module. De fleste talesynteseprogrammer ignorerer slik formatering, delvis fordi brukere ofte overkjører dette med sine egne preferanser (gjennom innstillingene i deres skjerm- eller dokumentlesere). Dette ble bekreftet av prosjektmedarbeider i samtale med forskere og studenter som bruker talesyntese daglig. Det må allikevel nevnes at tilpasningen av stilark for å støtte talesyntese via CSS Speech Module, selv om det for øyeblikket ikke er i bruk, hindrer ikke de brukere som ikke ønsker å bruke slik tilpasning.

Pga fordelene med EPUB-formatet som er nevnt ovenfor, bestemte vi oss å jobbe videre med EPUB. I tilfelle det blir mer utbredt støtte for CSS i flere talesynteseverktøy i fremtiden, bør stilark-tilnærmingen prøves ut igjen.

Andre fase: Universelt utformet EPUB

I arbeidet med EPUB, var vår hovedutfordring å utvikle dokumentmalen til Word som samtidig tilfredsstilte krav til universell utforming, tidsskriftets preferanser og som lot seg automatisk konvertere til EPUB. Forfattere kan benytte seg av Word-malen når de begynner å skrive manuskripter – alternativt kan malen tilknyttes til et eksisterende Word-dokument via Utvikler-funksjon i Word.

For konverteringen fra Word til EPUB brukte vi [Pandoc](#), et åpen-kilde kommandolinjeredskap som kan konvertere universelt mellom alle dokumentfiltyper.

Den første metoden vi prøvde ut var å konvertere direkte fra Word til EPUB. Utfordringen var at Word-formatet, som er mest brukt av forfattere og tidsskrifter, ikke inneholder nødvendig metadata, og dermed mangler EPUB-filen dette. Metadata kan legges til i EPUB manuelt eller med et eksternt verktøy. Det er gratis verktøy tilgjengelig på nettet som gjør det mulig å manuelt legge til metadata i EPUB-filer: vi benyttet programmet "[Calibre](#)" før vi innførte den overnevnte arbeidsflyten. Videre kan mer teknisk kyndige brukere legge til metadata gjennom egen programvare, eller direkte legge til tags og informasjon i EPUB-filens underfil, "content.opf". Fordelen med denne metoden er at den bevarer det samme visuelle uttrykket som i Word-malen, siden EPUB har tilsvarende formateringsmuligheter.

Vi ønsket imidlertid å unngå etterbehandling til den grad det var mulig – slik at redaktørene potensielt kunne klare å konvertere selv (vi kan ikke forutsette at redaktører har kompetanse til å redigere EPUB-filer). Derfor prøvde vi ut en metode der Word konverteres først til Journal Article Tag Suite (JATS) XML og så til EPUB.

[JATS](#)-formatet er et NISO-standard og som begynner å bli mer og mer etterspurt i vitenskapelig publisering pga formatets universell utforming, interoperabilitet og muligheter for langtidsbevaring. Standarden anbefales også av finansierer (f.eks Plan S). I Septentrio er det foreløpig ett tidsskrift, [Ottar](#), som prøver ut XML-flyten som et forsøk å gjøre tidsskriftet tilpasset for lesning på flere digitale flater (pc, nettbrett, mobil).

Med tanke på behovene i talesyntese- og Ottar-prosjektet, utviklet Septentrio-teamet et Pandoc-script som hentet nødvendige metadata fra OJS i konverteringsprosessen fra Word til JATS, plasserte metadata i JATS XML-filen, og deretter (som en talesyntese-spesifikk del) konverterte fra JATS XML til EPUB-filen, som dermed inneholder korrekte metadata og krever ingen etterbehandling. En ulempe med denne metoden er at noe av det visuelle uttrykket går tapt, fordi JATS har færre

formateringsmuligheter enn Word. Det visuelle uttrykket var dog ikke et prioritert for talesyntese-prosjektet – vår beslutning var at andre formater (f.eks PDF eller HTML) kan ev. brukes for et tidsskrift som ønsker å fremheve sitt grafiske uttrykk.

Pandoc-konverteringsverktøyet ble integrert i en [online app](#) (tilgjengelig via en webside-grensesnitt), der redaktører kan laste opp et Word-manuskript og laste ned ferdig konverterte filer.

Prosjektmedarbeider har utarbeidet en veileder til redaktører, som skildrer prosessen, steg for steg.

Pandoc-scriptet var utviklet for OJS 3.3, men i mellomtiden har det blitt lansert en ny versjon av OJS, 3.4. OJS 3.4 ser ut til å kreve færre tilpasninger i Pandoc-scriptet, som betyr at scriptet kan lettere tas i bruk av andre. Septentrio-teamet planlegger å tilpasse scriptet for OJS 3.4 og publisere scriptet på Github ila høsten 2023. Github-pakken kommer til å inneholde selve scriptet, nødvendig dokumentasjon av programvaren, og i tillegg utkast til en redaktørveileder (som en publiseringstjeneste kan tilpasse ved å legge til tjenestespesifikk informasjon) og eksempel i Word.

For endelig EPUB-filer, ønsket vi at de ikke bare skulle være nedlastbare, men også kunne vises direkte på tidsskriftssiden (for å bli opplest med egen skjermleser). Det finnes flere ulike programutvidelser for OJS som lar brukere lese EPUB-filer direkte, uten å bruke egen programvare. Standardvalget i OJS er for tiden "Bibi EPUB-viewer", men dette så ikke ut til å fungere godt, så vi valgte å installere [epubJSViewer](#). Eksempel på EPUB-fil som blir publisert i vår arbeidsflyt er en artikkel i *Nordic Perspectives on Open Science*: <https://doi.org/10.7557/11.6665>.

Videreføring og anbefalinger

I denne seksjonen summerer vi opp den endelige arbeidsflyten som prosjektet landet på etter intern testing, og presenterer anbefalinger til andre aktører innenfor universitetspublisering som ønsker å benytte seg av denne arbeidsflyten.

Forarbeid

Det må utarbeides dokumentmaler som tilfredsstiller kravene til tidsskriftet og til universell utforming, herunder valg av stil med høy kontrast og tydelige ordbilder, samt instruksjon til forfattere om å legge til alternativ tekst til bilder, figurer og tabeller (alternativ tekst beskriver på en tydelig måte for synshemmede hva disse elementene formidler). Det er viktig at forfattere blir instruert å bruke stiler i dokumentmalen isf å formatere manuelt. Manuelle justeringer, for eksempel å øke fontstørrelsen i stedet for å velge "overskrift 2", vil ikke fanges opp under automatisk konvertering, og fungerer heller ikke optimalt med talesynteseprogramvare, som bruker "tabbing" for å navigere fra overskrift til overskrift gjennom dokumentet.

Derneft må denne dokumentmalen testes mot Pandoc, hvor man har flere variabler som kan justeres etter ønske. Blant annet kan man lage et stilark (CSS) som brukes for alle EPUB-filene, og hvor man har tilpasninger for tilgjengelighet og eventuelt talemøduler. W3C har [EPUB og CSS eksempler](#) utformet for tilgjengelighet, som man kan laste ned og ta i bruk. I dette steget må man regne med litt feilsøking, og man sparer mye tid hvis man har et testdokument med mange mulige feilkilder (tabeller, figurer og lignende). Et annet viktig tips er å bruke [EPUBcheck](#), et valideringsverktøy fra W3C. Man burde også forsøke ulike tilleggspogram i OJS, dersom dokumentet ikke vises slik man ønsker, som første steg i feilsøking – blant annet kan man prøve en tidligere versjon av samme app.

Her må man også ta stilling til hvordan man ønsker å føre inn korrekt metadata i EPUB-filen. Vårt foretrukne alternativ er å bruke en tilgjengelig JATS XML-fil som inneholder metadata anbefalt av Dublin Core – metadataene settes inn ved hjelp av et tilpasset Pandoc-OJS-konverteringsscript.

Det siste steget er å aktivere et program for å vise EPUB-filer i OJS, slik at brukere som ønsker talesyntese kan velge mellom å laste ned filen for bruk i dokumentlesere eller lese den direkte på siden med skjermlesere.

Rutiner

Forfattere er forventet å levere sitt endelige utkast i henhold til dokumentmalen som er tilgjengelig på tidsskriftets hjemmeside. Både forfatter og redaktør har ansvar for å sikre at korrekt metadata er lagt til i OJS. Før konverteringen bør redaktøren også kontrollere at bildene i Word-filen har alternativ tekst for synshemmede og sikre at formateringen er korrekt. Typiske problemer kan være manglende stilformatering på overskrifter eller feilformaterte kildehenvisninger, ofte fordi forfatterne har formatert manuelt, i stedet for å bruke stiler, eller der hvor programmer som Endnote eller Zotero har gjort automatiske endringer. Hvis det oppdages feil, bør redaktøren eller forfatteren rette disse i Word-dokumentet. Når dette er gjort, kan redaktøren laste opp Word-filen på vår konverteringsside, hvorpå en Zip-fil med manuskriptet i formatene PDF, EPUB, JATS XML og HTML blir generert. EPUB-filen lastes opp som en "galley" i OJS.

Før publisering er det mulig å forhåndsvisne EPUB-filen. Hvis det oppdages formateringsfeil, må disse rettes i Word-filen og konverteringen må gjøres på nytt.

Oppsummering

Det er et bredt spekter av både åpne og proprietære programvarer for talesyntese tilgjengelig, og vi har konkludert med at EPUB vil nå flest brukere. På nåværende tidspunkt er det stor usikkerhet om hva som vil bli den fremtidige standarden, spesielt siden talesyntese stadig forbedres ved hjelp av kunstig intelligens. Denne usikkerheten indikerer at man ikke bør binde seg til mer begrensede formater, som SSML. I stedet kan man holde seg til PDF som låst format for å bevare et bestemt

uttrykk, og EPUB som åpent format for interaktivitet, lesbarhet og kompatibilitet med talesyntese. EPUB har klare fordeler, som støtte for paginering (noe som er viktig for forskere og studenter som må referere til permanente sidetall), og mulighet for detaljerte metadata, som er sentrale for indeksering, reproduserbarhet og langsiktig bevaring. Man må være klar over at selv om selve konverteringen kan gjøres av tidsskriftsredaktører uten mye teknisk kompetanse, må en publiseringstjeneste som ønsker å ta denne løsningen i bruk ha tilgang til en del teknisk kompetanse: skriptet for Pandoc-konvertering må tilpasses slik at det er mulig å hente ut metadata fra den relevante OJS-instansen, det er av og til behov for feilsøking i EPUB-filer (som gjøres i en EPUB-validator), og oppgraderinger av de involverte programvarene kan også skape behov for tilpasninger i skriptet og arbeidsflyten.

Likevel kan vi påstå at vi har kommet frem til en løsning som er realistisk og fremtidsrettet. Det har vært svært mye feilsøking involvert i prosessen, men andre brukere kan benytte seg av vår løsning, og våre vurderinger, uten å måtte starte på bar bakke.

Prosjektet ble presentert i en tidlig fase på et seminar som hadde universell utforming som fokus og var rettet mot de norske OJS-baserte institusjonelle publiseringsplattformer. Seminaret ble avholdt 22.-23. november 2021 og hadde flere foredrag.

I arbeidet med prosjektet var vurderingene våre informert av andre brukere, spesielt fra [PKP forumet](#) (diskusjonsplattform for Public Knowledge Project, utvikleren av OJS) og [Pandoc-forumet](#). Vi har også delt våre egne vurderinger og løsninger i disse fora. Videre utvekslet vi erfaringer rundt XML-baserte publiseringsløsninger innenfor det norske OJS-miljøet.

Et abstrakt for en presentasjon ble sendt inn til et satellittmøte ved konferansen for IFLA (International Federation of Library Associations and Institutions). Satellittmøtet hadde universell utforming og skulle avholdes 18.-19. august 2023 i Rotterdam, Nederland. Abstraktet ble dessverre ikke akseptert. Det blir presentert en poster om prosjektets resultater og erfaringer på den 18. [Munin-konferansen](#) 8.-10. november 2023 i Tromsø.