

Diplomatarium Menotae

Ivar Berg, 10.08.2018

Nokre merknader til koding av diplom etter Menota-standardten og inkludering i Menota-arkivet.

1. Vising: Heilt banalt: Eit *Diplomatarium Menotae* (*Menota* føyer seg naturleg inn i første dekl. med GEN.SG -ae) må bli eit subkorpus i Menota-arkivet, elles blir katalogen heilt uoversiktleg. Det bør nok vera mogleg med ei vising der ein kan bla og sortere på år og stad, men det treng ikkje vera framsida.

Eg ønskjer at dei mest sentrale opplysingane i headeren visest på skjermbiletet, slik at ein ikkje må leite i XML-fila etter dei. Eg vil altså ha ei vising som liknar på *Diplomatarium Norvegicum* (DN) på nett og som minimum viser datering og heimfesting, helst òg referanse til DN og RN (*Regesta Norvegica*) som hyperlenkjer til nettutgåvene. Oppsettet kan vera som i DN, slik at denne informasjonen står ovanfor teksten.

2. Header

Må inngå: Referanse til DN og RN (band og nummer), dato og stad. Den fremste føremonen med diploma som språkhistoriske kjelder er at dei er originalar som normalt er tid- og heimfeste, og derfor bør denne informasjonen vera lett tilgjengeleg utan å slå opp andre stader. Opplysingar om tid og stad kan byggje på RN med same kodar for uvisse (t.d. [usikker dato]).

Noko av dette kan leggjast i <msDesc>, men for at det skal visast i Corpuscle, bør det kanskje heller leggjast inn som ei eiga blokk (<div>) heilt først i teksten? Da tenkjer eg på dei opplysingane vi (eg) gjerne vil ha over sjølve tekstblokka i manuskriptvisinga. Den tekniske løysinga må eg overlata til kunnige folk.

Stad: Det er svært varierende lokalitetar som blir bruka i utgåvene. Her bør ein derfor kunne oppgje noko svært lokalt, t.d. gardsnamn, og noko lettare gjenkjenneleg, t.d. kommune. Dette bør utan veldig store vanskar kunne knytast til koordinatar og ei kartlenkje.

Språk: Ikkje oppgjeve i DN/RN, men bør vera med. Først og fremst må det vera mogleg å *søkje* på språk (eller ekskludere frå søket). Det kan bli problematisk for diplom i seinmellomalderen, men er heilt greitt i alle fall til eit godt stykke ut på 1400-talet. For seinmellomalderen bør «skandinavisk» vera eit definert språk som overkategori for dansk, norsk og svensk. (Kanskje to kategoriar, «hovudspråk» og «sidespråk»? Mange diplom på norsk inneheld t.d. latinsk helsing og datering.)

Det er viktig at systemet gjer det mogleg å søkje på all informasjonen i headeren. Her er diploma ulike dei andre tekstane i Menota i det at dei er korte tekstar, men mange. I forskingssamanheng er det ofte tenleg å sortere ut tekstar frå ein bestemd periode, eit bestemt geografisk område e.l.

3. Koding/annotasjon

Minimum diplomatarisk transkripsjon. Lemmatisering bør vera eit mål for alle publiserte tekstar, medan meir utførleg morfologisk og syntaktisk annotasjon er eit lengre lerret å bleikje.

Det bør òg vera eit mål å lemmatisere person- og stadnamn. Er det kanskje mogleg å innføre person-id-ar, der ein kan slå fast at same person er omtala i fleire tekstar? Her blir det naturlegvis tvilstilfelle, men særleg for dei høgge standa er det ofte opplagt kven det er snakk om. Dette vil opne for interessante studiar av skrivemiljø og ulike former for variasjon (individuell, sosial).